THE
PHYSICAL REVIEW

A journal of experimental and theoretical physics established by E. L. Nichols in 1893

SECOND SERIES, VOL. 97, NO. 2 JANUARY 15, 1955

Departure from Paschen's Law of Breakdown in Gases

W. S. BOYLE AND P. KISLIUK
Bell Telephone Laboratories, Inc., Murray Hill, New Jersey
(Received September 28, 1954)

The failure of Paschen's law of electrical breakdown in gases at both high pressures and extremely small electrode separations is explained by a single breakdown mechanism. The breakdown field in each of these cases is sufficiently great to draw measurable field-emission current from the cathode, which produces a relatively small number of ions. The space-charge field of these ions is great enough to increase the field-emission current appreciably even when the ratio of ion current to electron current is less than one percent. As more ionic space charge is produced, each ion becomes more effective in enhancing the electron current until the breakdown condition is attained. An expression is derived for the yield of electrons per positive ion as a function of the applied field. This expression is shown to be in quantitative agreement with values of γ derived from published data on breakdown voltages at high pressures. It is shown also that the same process explains breakdown at extremely small separations below the "minimum sparking potential."

INTRODUCTION

PASCHEN'S law[1] of electrical breakdown in gases states that the breakdown voltage, V_B, is a function of the pressure, p, (or, more precisely, the density, ρ) and the electrode separation, d, only in the combination pd $[V_B=f(pd)]$. Since the earliest investigations of gas breakdown a considerable literature has accumulated on the failure of Paschen's law, expecially at high pressures.[2-4] Typical behavior showing departure from this law for plane parallel electrodes at high pressures is shown in Fig. 1, where with increasing distance the breakdown voltage increases linearly, while with increasing pressure the voltage starts to increase linearly but the slope falls off at high pressures.

Departure from Paschen's law occurs also for breakdown between closely spaced electrodes at atmospheric pressure with potential differences below the "minimum sparking potential."[5,6] In the case of very small separations, breakdown occurs at potentials for which breakdown does not occur at any pressure if the electrodes are fixed at some larger spacing. The breakdown voltage is thus clearly not a function of the product pd alone. The behavior of the breakdown voltage at small spacings is shown in Fig. 2.

The deviations from Paschen's law occur for just those experimental conditions for which the multiplication of electrons by ionization in the gas is very small. For these conditions the streamer mechanism[7] and the usual Townsend process involving values of the gamma coefficient of the order of magnitude of those usually measured[8] fail to account for the current multiplication necessary to attain breakdown. The theory developed here accounts for breakdown under both of these conditions for which Paschen's law fails. This theory was suggested for low-voltage breakdown by Germer and Haworth[5,9] and for breakdown at high pressures first by Zeier[10] and later by other workers,

[1] F. Paschen, Ann. Phys. u. Chem. **37**, 69 (1889).
[2] A. H. Howell, Trans. Am. Inst. Elec. Engrs. **58**, 193 (1939).
[3] A. Boulloud, Ann. phys. **8**, 909 (1953).
[4] D. R. Young, J. Appl. Phys. **21**, 222 (1950).
[5] L. H. Germer and F. E. Haworth, Phys. Rev. **73**, 1121 (1948).
[6] P. Kisliuk, J. Appl. Phys. **25**, 897 (1954).

[7] L. B. Loeb and J. M. Meek, *The Mechanism of the Electric Spark* (University Press, Stanford, 1941).
[8] M. J. Druyvesteyn and F. M. Penning, Revs. Modern Phys. **12**, 87 (1940).
[9] L. H. Germer and F. E. Haworth, J. Appl. Phys. **20**, 1085 (1949). This mechanism for breakdowns in air at extremely short distances has been considered also in reference 6 and by M. M. Atalla, Bell System Tech. J. (to be published).
[10] O. Zeier, Ann. Physik **14**, 415 (1932). Several authors have subsequently suggested this mechanism for breakdown at high pressures, notably Trump, Cloud, Mann, and Hanson, Elec. Eng. **69**, 961 (1950); F. L. Jones and C. G. Morgan, Phys. Rev. **82**, 920 (1951); A. Boulloud, Compt. rend. **233**, 932 (1951) and reference 4. None of these authors develops the ideas further, and some of them suggest alternatives. These alternatives are largely discredited in reference 3, where the author finds support for this mechanism by a process of elimination.

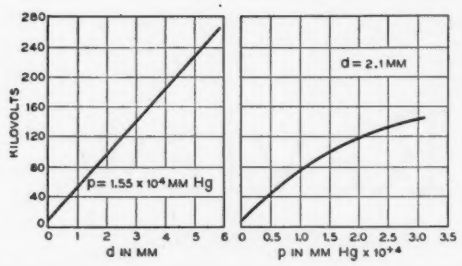


FIG. 1. Breakdown voltage plotted against distance for constant pressure, and against pressure at constant distance, showing the failure of Paschen's law at high pressure (data from reference 2).

but none of these developed the consequences of the theory beyond its initial conception.

In brief, the theory is as follows. When gas multiplication is small, the electric field becomes sufficiently large before breakdown to draw field-emission currents from small irregularities on the cathode. The field-emission electrons produce a small number of ions which, by their space charge, increase the field at the cathode. Because of the steep increase in current with field, this process is able to build up to breakdown even when the probability of any particular electron having an ionizing collision is extremely small.

Paschen's law is based on the observation that, over a large range of pressures and electrode separations, the probability of ionization per collision in the gas and the probability of the production of electrons by ions by a secondary process are both dependent on the average kinetic energy of the electrons and ions, and therefore[11] on E/p. The proposed process of electron emission due to the combined applied and ionic-space-charge fields depends primarily on E rather than E/p, leading directly to deviations from Paschen's law.

It is shown below that the process, whereby field emission is enhanced by the increased field due to approaching positive ions, will ultimately become unstable and breakdown will occur. In this paper this theory is developed in quantitative form and an explicit expression is obtained for γ', the effective yield of electrons per ion at breakdown. This expression may be compared with experimental data obtained from the literature on electrical breakdown in air at high pressures using the published values of η, the ionization coefficient per volt along an electron path.[12] The experimental values of γ' are found to vary with applied field at breakdown in the way predicted by theory, confirming the proposed mechanism. The theory accounts also for low-voltage breakdown across very short gaps in air at atmospheric pressure and in vacuum, although in these cases experimental data are not available for a quantitative check.

[11] L. B. Loeb, *Fundamental Processes of Electrical Discharge in Gases* (John Wiley and Sons, Inc., New York, 1939), pp. 410–412.
[12] F. H. Sanders, Phys. Rev. 44, 1020 (1933).

THEORY

Consider the empirical form of the field emission equation:

$$j_E = A \exp(-B/E), \tag{1}$$

where j_E is the electron current density, E is the field at the cathode, and A and B are constants. Let $E = E_A + E^+$, where E_A is the applied field and E^+ the field due to the positive-ion space charge. Because of the steepness of the dependence of current density on field, $E^+ \ll E_A$ and by expansion we find

$$j_E \approx A \exp(-B/E_A) \exp(BE^+/E_A^2) = j_0 \exp(BE^+/E_A^2), \tag{2}$$

where $j_0 = A \exp(-B/E_A)$ is the current which would flow due to the applied field in the absence of space charge. Furthermore, to a good approximation,

$$E^+ = Cj^+, \tag{3}$$

where C is a constant involving the ionic mobility and the electrode microgeometry and j^+ is the positive-ion current density. Let

$$j^+ = Gj_E^n, \tag{4}$$

where G is a constant involving the electron and ion trajectories and n is to be evaluated in the later discussion. Then, $E^+ = CGj_E^n$, and Eq. (2) becomes

$$j_E = j_0 \exp(BCGj_E^n/E_A^2) = j_0 \exp(Mj_E^n), \tag{5}$$

where, since E_A is a very weak function of j_E, the coefficients of j_E^n have been included in a single constant, $M = BCG/E_A^2$.

The breakdown condition for this process can be determined by examining Eq. (5). Figure 3 is a plot of the left- and right-hand sides of Eq. (5) for the particular case $n = 2$. For small j_0 there is a stable solution, whereas for large j_0 there is no solution and the breakdown condition has been exceeded; breakdown

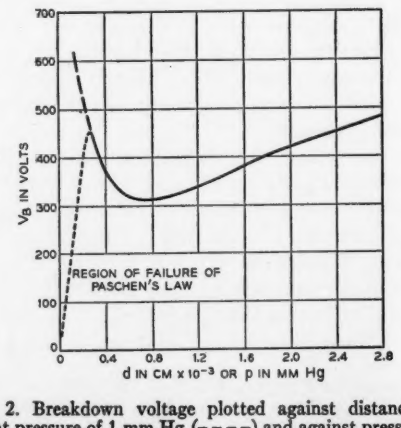


FIG. 2. Breakdown voltage plotted against distance at a constant pressure of 1 mm Hg (– – – –) and against pressure at a constant distance of 1 cm (— — —), to show the failure of Paschen's law at short distances. In the solid portion (——) the two curves coincide and Paschen's law is valid.

occurs when the curves are tangent. Thus, the critical condition for breakdown is $(\partial/\partial j_E)[j_0 \exp(Mj_E{}^n)]=1$, which with Eq. (5) can be written $M=1/nj_E{}^n$, or

$$j_E=j_0 \exp(1/n). \tag{6}$$

Expressed in words, Eq. (6) states that breakdown occurs when the applied field is such that the field-emission current is increased by space charge by just the factor $\exp(1/n)$.

Let us now define the effective γ as the ratio of the field-emission electron current density to the incoming positive-ion current density,

$$\gamma=j_E/j^+=1/Gj_E{}^{n-1}=BC/E_A{}^2Mj_E{}^{n-1}, \tag{7}$$

by Eq. (4). From Eq. (6) this becomes at breakdown,

$$\gamma'=BCnj_E'/E_A{}'^2 = \{BCn[\exp(1/n)]/E_A{}'^2\}A \exp(-B/E_A'), \tag{8}$$

where γ', j_E', j_0', and E_A' are the values of these variables just at breakdown. Since the exponential dependence is strong enough to mask the variation in $1/E_A{}'^2$, γ' is given to a good approximation by,

$$\gamma'=K \exp(-B/E_A'), \tag{9}$$

where K is a constant and B the same constant that appears in the empirical expression for field emission, Eq. (1). Note that while K depends on n, the form of Eq. (9) is independent of n. This theoretical relation is a direct consequence of the theory of field emission breakdown. Confirmation of the theory will be obtained if there is a satisfactory agreement between Eq. (9) and published experimental data on breakdown at high pressures.

The usual formulation of the breakdown condition in terms of ionization by electrons and a secondary process of electron production is

$$\gamma'(e^{\eta'V'}-1)=1, \tag{10}$$

where $(e^{\eta'V'}-1)$ is the number of ionizing collisions resulting from the emission from the cathode of one electron, V' is the breakdown voltage, and η' is the number of ionizing collisions per volt of potential difference at breakdown. The quantity η is a strong function of E/p and values of η as a function of E/p are available in the literature for some gases. Thus if the pressure, p, the breakdown voltage, V', and the separation are known, γ' can be computed from Eq. (10). If breakdown data for a range of pressures are available, the experimental dependence of γ' on E_A' will be determined.

The comparison between theory and experiment can be made reliably only for the case of breakdown at high pressure. For breakdown at low voltage the material evaporated by pre-breakdown field emission currents may seriously affect the pressure which is therefore not sufficiently well known to compute γ'. Experimental data for three different conditions are considered below: breakdown at high pressure, low-voltage breakdown at atmospheric pressure, and low-voltage breakdown in high vacuum.

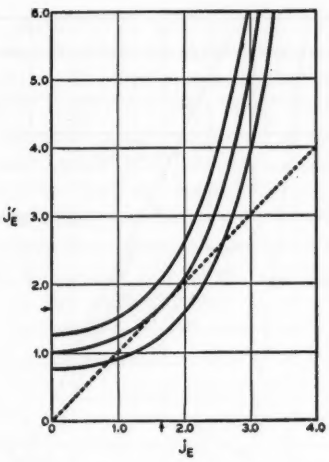


FIG. 3. Plot of $j_E'=j_E$(- - -), and $j_E'=j_0 \exp(Mj_E{}^3)$(——) for various values of j_0, in arbitrary units selected so that $j_0=1$ at breakdown.

Breakdown at High Pressure

At high pressure the exponent n in Eq. (4) is unity. The coefficients C and G are not strictly constants but their variations are second-order effects which leave the exponential form of Eq. (9) essentially unaffected. The coefficient B of Eqs. (1) and (9) can be obtained in terms of the apparent applied field from the Fowler Nordheim equation.[13] The expression is

$$B=6.85\times10^7\Phi^{\frac{3}{2}}/\beta, \tag{11}$$

where Φ is the work function and β the maximum field multiplication due to the surface irregularity giving rise to the emission. $\Phi^{\frac{3}{2}}$ is of the order of magnitude of 10, while β is of the order of 50, whence $B\approx10^7$ volts/cm.

Plots of $\log\gamma'$ against $1/E_A'$ from experimental data taken in air are given in Fig. 4. It can be seen that the points fall on satisfactory straight lines. In Table I are values of B obtained from these plots. These are of the right order of magnitude.

It would be desirable to find an experimental check for gases in addition to air. There are very few gases, however, for which data are available for both of the required quantities, namely η at small values of E/p and the breakdown voltage between plane parallel electrodes at high pressure. Only N_2, H_2, CO_2, and air appear to be possibilities at present. Except for air, the agreement with Eq. (9) is poor.

This may be due, in part, to the effect of small concentrations of impurities which become important at high pressures, especially for gases of high ionization

[13] R. H. Fowler and L. Nordheim, Proc. Roy. Soc. (London) **A119**, 173 (1928).

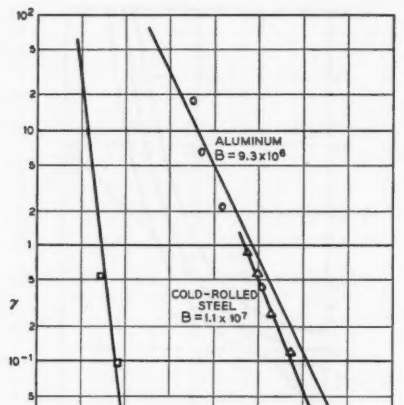


FIG. 4. γ computed from breakdown voltages in air plotted against the inverse of the apparent breakdown field $(1/E_A)$, from the data of reference 3 (stainless steel and aluminum) and reference 2 (cold rolled steel), using η values from reference 12.

potential. In other words, data on η taken for nearly pure gases at a particular value of E/p at low pressure may not be applicable to the same gases at the same value of E/p at much higher pressure. Furthermore there is direct evidence for CO_2 that η may not be a function of E/p alone at pressures approaching the condensation point.[4] Thus an experimental check of Eq. (9) is probably reliable only for a mixture including a considerable concentration of impurities, and air is the only such mixture for which we have data.

The proposed mechanism seems to be justified by the satisfactory fit of the straight lines in Fig. 4 and by the agreement of the numbers in Table I with the estimated value of B. A complete check, including computation of the constant K of Eq. (9), may be quite difficult, especially if there are complications due to electron attachment, ionization by metastables, etc., but fortunately knowledge of this constant is not necessary. It would be useful, however, in computing maximum pre-breakdown current densities and accurate breakdown fields.

A number of experimental observations can be explained by the proposed mechanism. The breakdown field between plane parallel electrodes at high pressure is expected to depend strongly on the work function and roughness of the cathode surface. This is justified by a great many investigations.[2-4] An observation by Howell[2] that anode roughness also affects the breakdown voltage may be discounted because his treatment of the anode must have left loose particles which, in the presence of the high field, will be exchanged between electrodes.

Young[3] has observed that deviations from Paschen's law occur at lower pressure for smaller electrode spacings. This is to be expected because the decreased gas multiplication at smaller spacings allow the breakdown fields to increase to the values necessary for field emission at lower pressures.

Low Voltage Breakdown in Air

Assume for the moment, that the pressure between the electrodes is unknown, either because of material evaporated by pre-breakdown currents or because, in the case of closing electrodes, air may be compressed between them. The ionization coefficient, η, cannot have a value more than its maximum which is ≈ 0.015 for air.[8] Since breakdowns have been observed as low as 30 volts,[7] breakdown is possible when $(e^{\gamma V}-1)$ is less than 0.65, whence, for breakdown, [by Eq. (15)] $\gamma > 1.5$. This is very large, typical values observed in glow discharges[8] being of the order of 0.01. If the pressure is atmospheric, E/p is much greater than the value corresponding to a maximum η for the observed breakdown fields, so that η is well below its maximum, and γ may well be in excess of four electrons/ion. This is consistent with the observed very short times of current buildup in these breakdowns.[14,6]

The observed breakdown fields are of the order necessary to draw measurable field emission currents when the same electrodes are placed in vacuum. Thus the situation is proper for a small ionic space charge to increase the current greatly as described previously.

Experiments have been conducted to observe the pre-breakdown currents when electrodes are closed in air with potentials less than the minimum sparking potential ($\sim$300 v). When the potential difference was above 180 volts, these currents were too small to be observed on the equipment used (<3 μa). For the lowest potential differences at which breakdown could be observed ($\sim$30 volts), the pre-breakdown currents increased to about 20 microamperes. This increase in pre-breakdown current is undoubtedly a result of the higher breakdown field needed, according to Eq. (9), to obtain the required high value of γ.

The value of n in Eq. (5) is uncertain for low-voltage air breakdown because it is not known whether

TABLE I. Experimental values of the slope of the logγ *versus* $1/E$ curves.

	B(volts/cm)
Cold rolled steel[a]	1.1×10^7
Stainless steel[b]	2.6×10^7
Aluminum[b]	9.3×10^6

[a] Data from reference 2.
[b] Data from reference 3.

[14] L. H. Germer and J. L. Smith, J. Appl. Phys. **23**, 553 (1952).

the pressure between the electrodes depends on the pre-breakdown current. It is, in any case, between 1 and 3.

Low-Voltage Breakdown Between Closely Spaced Electrodes in High Vacuum

In this case the pre-breakdown currents may become much larger than in air, amounting to several milliamperes for breakdown voltages of the order of one kilovolt. The power dissipated on the anode is then sufficient to evaporate metal, breakdown occurring in the resulting vapor.[15] The electron current density, in addition to being the primary means of ionization, affects the positive ion current density by producing the vapor, the pressure of which determines η and the ionic mobility. The power of j_E occurring in Eq. (5) is thus between two and three depending on whether the breakdown pressure is sufficient to affect the velocity of the ions.

Examination of pulsed field-emission currents with uniformly increasing applied voltage up to the point of breakdown fail to show any observable evidence of enhancement of the current directly by ionization in the gas.[15] This is consistent with the field-emission mechanism of breakdown because a very small fraction of ionic current is sufficient to increase the field-emission current to the point where Eq. (6) is satisfied.

CONCLUSION

A quantitative treatment of the role played by field emission in the electrical breakdown of gases has been given. In particular, a new criterion for electrical breakdown has been formulated which appears to have wide application. The regenerative field-emission breakdown process applies whenever large fields sufficient to draw field-emission currents appear at the cathode. This situation arises when, either because of high pressure or close electrode separation, the probability of an electron making an ionizing collision is small.

The deviation from Paschen's law which occurs for all gases at sufficiently high pressure is satisfactorily explained. In the particular case of air, where the ionization coefficient is known over the appropriate range of E/p, the theory is in very good quantitative agreement with the form of the breakdown curve at high pressures.

For the case of closely spaced electrodes in air, it is known that breakdown may occur at a potential difference far below the minimum sparking potential determined by a Townsend mechanism involving the usual measured values of γ. It has been shown that this limitation does not apply for the field-emission process, however, and that breakdown will always occur as long as the potential difference between the electrodes is sufficient to give at least some ionization of the gas and the electrode separation is sufficiently small to lead to a certain critical field emission current.

It should be noted finally that this breakdown criterion should apply equally well to liquids of high dielectric strength where once again electrical breakdown is preceded by the flow of field-emission currents.

[15] Boyle, Kisliuk, and Germer, J. Appl. Phys. (to be published).

PHYSICAL REVIEW VOLUME 97, NUMBER 2 JANUARY 15, 1955

Nuclear Susceptibility of Liquid and Solid He^3

P. J. PRICE
I. B. M. Watson Scientific Computing Laboratory, Columbia University, New York, New York
(Received September 1, 1954)

An interpretation of the observed deviation from Curie's law of the nuclear paramagnetic susceptibility of liquid helium He^3, based on a "cell" type of model of the liquid, is presented. The theory is able to account qualitatively, but not precisely, for the deviation. The reasons for its imprecision are discussed. The bearing of the point of view of the theory on the paramagnetism of the solid is considered: it is concluded that there should be no deviation from Curie's law at the lower temperatures, but that a small deviation near the rising part of the melting curve is quite possible.

LIQUID helium He^3 has a magnetic susceptibility due partly to the diamagnetism of the atomic electron shells and partly to the paramagnetism of the orientation distribution of nuclear magnetic moments. The latter contribution has recently been measured separately, by the method of resonance absorption, by Fairbank *et al.*[1,2] This paramagnetic susceptibility (which we shall denote by χ, since there will be no need of a symbol for the diamagnetism) is of interest in the theory of quantum liquids, because its magnitude is a direct measure of the influence of quantum statistics. The actual magnetic interaction energy of neighboring atoms is very small compared with kT at any attainable temperature; and hence, if Boltzmann statistics applied, the value of χ would be

$$\chi_B = N\mu^2/kT \text{ per mole}, \quad (1)$$

[1] Fairbank, Ard, Dehmelt, Gordy, and Williams, Phys. Rev. 92, 208 (1953).
[2] Fairbank, Ard, and Walters, Phys. Rev. 95, 566 (1954).

where $\mu=1.07\times10^{-23}$ erg/gauss is the magnetic moment of a He^3 nucleus (the spin moment of the nucleus being $\hbar/2$). The ratio

$$F(T)\equiv\chi/\chi_B \tag{2}$$

thus measures the effect of Fermi-Dirac statistics, in restricting the wave functions (including spin) of the system to the set antisymmetric in the atomic variables. The conclusion of the spin resonance experiments (reference 2, Fig. 1: the quantity $\chi T/C$ graphed there is our F) is that F falls off appreciably from unity at 1°, and by about 0.2° has fallen to 0.6 and is decreasing quite rapidly with decreasing temperature. Such a deviation from Curie's law is of course required[3] by the "ideal Fermi gas" model[4] of liquid He^3; but the quantitative prediction of this model is a decline of F at considerably higher temperatures—e.g., to 0.8 at 5°—than was found. It is a strange-seeming property of this model, however, that the experimental values of $F(T)$ can be made to fit the theoretical curve quite well if in the latter a degeneracy temperature of 0.45° is arbitrarily substituted for the value of 5° calculated from the density.[2] It is possible that a convincing explanation of this shift will be provided by a future elaboration of the ideal gas model. The physical basis of this model is, however, questionable; and it is the main purpose of this paper to show how the deviation from Curie's law may be approximately accounted for by a rough but basically more realistic treatment.

The starting point is an exact formula for the molar paramagnetic susceptibility:

$$\chi=\langle M_z^2\rangle/kT, \tag{3}$$

where M_z is a component of the total magnetic moment $\mathbf{M}$ due to all N nuclear moments and $\langle\,\rangle$ denotes the average in zero external field over the canonical distribution at temperature T. This formula (3) is proved in the appendix. If μz_n is the z component of the magnetic moment of the nth nucleus (the eigenvalues of z_n being ±1), then

$$M_z^2=\mu^2[\sum_{n=1}^{N} z_n^2+\sum_{m\neq n}\sum z_m z_n],$$

and hence

$$\langle M_z^2\rangle=N\mu^2[1+2(N-1)\langle z_1z_2\rangle]. \tag{4}$$

In the high-temperature Boltzmann-statistics limit the nuclear spins are uncorrelated, and $\langle z_1z_2\rangle=0$. Substituting (4) in (3), we then retrieve (1). In general, we have

$$F=\langle(\textstyle\sum_n z_n)^2\rangle/N=1+2(N-1)\langle z_1z_2\rangle. \tag{5}$$

This result (5) is exact. The spin correlation $\langle z_1z_2\rangle$ is of order $1/N$ because—in configuration terms—at any instant a fraction of order $1/N$ only of the atoms is near enough to any one atom for its nuclear spin state to be influenced by theirs.

Equation (5) enables us to deduce the susceptibility directly from a suitable approximate model of the liquid state, rather than by attempting to calculate the perturbation of the model by a magnetic field. We shall show, in terms of a "cell" type of model,[5] how the deviation from Curie's law may be caused by the temporary confinement of pairs of atoms in spaces defined by their neighbors. In order to make a tractable calculation of this effect, we suppose that at any instant the atoms may be divided into $N/2$ pairs of neighbors so that the spins of each pair are mutually correlated but there is no net inter-pair correlation. The wave function is supposed to be approximately separable, so that each pair moves in the "cell" field of the average potential of interaction with the surrounding atoms, together with the mutual interaction potential of the pair. If $\langle\,\rangle_c$ denotes the average only over the time when a pair of atoms are together in a cell, then (remembering not to "count pairs twice") we have

$$F-1=\langle z_1z_2\rangle_c. \tag{6}$$

The average in (6) is taken to be over the canonical distribution of the pair states. The wave functions of the latter factorize into orbital and spin functions, $\psi=\phi(\mathbf{r}_1,\mathbf{r}_2)S(z_1,z_2)$ and may be classified as *singlet*, for which $\psi=\psi_p=\phi^+S^-$ and *triplet*, for which $\psi=\psi_0=\phi^-S^+$, where the $+$ and $-$ denote symmetric and antisymmetric functions. For each ϕ^+ orbital there is one state, for which $z_1z_2=-1$, and for each ϕ^- orbital there are three states, for which the average of z_1z_2 is $+\frac{1}{3}$. Hence if

$$Z_p=\sum\exp(-E_p/kT),\quad Z_0=\sum\exp(-E_0/kT), \tag{7}$$

are respectively the separate "singlet" and "triplet" partition sums (reckoned without including the threefold spin degeneracy of the triplet states), E_p and E_0 being the energy eigenvalues, then the average of the spin correlation (6) over the canonical distribution is

$$\langle z_1z_2\rangle_c=(Z_0-Z_p)/(3Z_0+Z_p).$$

Substituting this result in (6), we have

$$F=4Z_0/(Z_p+3Z_0). \tag{8}$$

The variation of F with temperature thus depends on the E_p and E_0 spectra. A detailed calculation of the level scheme would not be justified here,[6] because of the essential crudeness of the model and the ambiguity in choice of the precise form of cage potential field. The essential features of the model, which determine its magnetic properties, are that (a) the pair interaction imposes a lower limit on the distance $|\mathbf{r}_1-\mathbf{r}_2|$ between

[3] L. Goldstein and M. Goldstein, J. Chem. Phys. **18**, 538 (1950).

[4] I. Pomeranchuk, J. Exptl. Theoret. Phys. (U.S.S.R.) **20**, 1919 (1950).

[5] J. E. Lennard-Jones and A. F. Devonshire, Proc. Roy. Soc. (London) **A163**, 53 (1937), Sec. 3.

[6] Such calculations have been made for a single atom in a cage, for liquid He^4, by F. London, Proc. Roy. Soc. (London) **A153**, 576 (1936), and by C. L. Pekeris, Phys. Rev. **79**, 884 (1950).

atoms; (b) the cage field confines the atomic positions $\mathbf{r}_1,\mathbf{r}_2$ separately to a small spherical volume. We could reasonably assume the effective potential to be otherwise constant, and work out the wave functions and energy levels for "two spheres in a well"; but this procedure would still be unjustifiably complicated. Instead we will simply set the interatomic distance $|\mathbf{r}_1-\mathbf{r}_2|$ equal to a fixed average value d. The energy eigenvalues are then just those of the rigid rotator:

$$\begin{cases} E_p=l(l+1)(\hbar^2/md^2); & l=0,2,4,\cdots \\ E_0=l(l+1)(\hbar^2/md^2); & l=1,3,5,\cdots \end{cases} \tag{9}$$

where m is the mass of a He^3 atom. For each l value there is an orbital degeneracy of $2l+1$ levels. Accordingly

$$\begin{cases} Z_p=\sum_{l\ \text{even}}(2l+1)\exp[-l(l+1)T_0/T], \\ Z_0=\sum_{l\ \text{odd}}(2l+1)\exp[-l(l+1)T_0/T], \end{cases} \tag{10}$$

where

$$T_0=\hbar^2/md^2k. \tag{11}$$

The partition sums (10) are those used in the theory of ortho- and para-hydrogen, and have already been tabulated.[7] Making use of these results, we obtain from (8) and (10) a theoretical $F(T/T_0)$.

Figure 1 shows the theoretical curve of $F(T)$ compared with the experimental results. The parameter T_0 of (10) has been arbitrarily chosen so that the two agree at $T=0.4°$. This requires that $T_0=0.39°$. Substitution of this value in (11) gives $d=6.4\times10^{-8}$ cm. To assess the plausibility of this latter value, we remember that it is that of the imaginary fixed distance between the centers of two helium atoms, when rotating freely about each other. The volume taken up by this motion is $\pi(d+a)^3/6$, where a is the effective closest approach distance between atoms: this should equal, approximately, the mean volume occupied by two atoms. Thus the value, d' say, of d arrived at on these considerations is given by

$$(d'+a)^3\simeq 12/\pi n, \tag{12}$$

where n is the number density of atoms in liquid He^3. Substituting $n=1.6\times10^{22}$ cm^{-3}, we obtain $d'+a\simeq6.2\times10^{-8}$ cm. Hence $d'\simeq4\times10^{-8}$ cm. Thus the theory given here is able roughly to account for the effect of quantum statistics on the susceptibility, down to a few tenths of a degree. The essential reason why it was not possible to derive the precise quantitative law lies in the pair approximation leading to (6). In actuality a given atom is surrounded by a number of nearest neighbors, and simultaneous correlation of its spin with each of theirs is to be expected.

[7] See J. E. Mayer and M. G. Mayer, *Statistical Mechanics* (John Wiley and Sons, Inc., New York, 1940), pp. 450–452. In their notation, $Z_p=Q_{jmo}$, $Z_0=Q_{jmo}$, $Q_{jmo}+Q_{jmo}=Q_{jmt}$, $T_0/T=\sigma$.

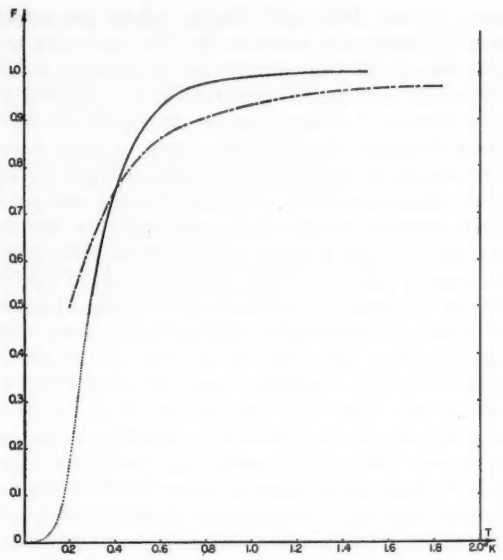


FIG. 1. The chain line represents the experimental results, and the full line the theoretical formula (8), (10) with $T_0=0.39°$. The dotted line is a schematic extension of the full one.

The agreement should get poorer at still lower temperatures. When $T\ll T_0$, the theory gives

$$F\simeq 12\exp(-2T_0/T). \tag{13}$$

According to (13), $F(T)$ and its derivatives vanish at $T=0$, as shown in the schematic extension (dotted) of the theoretical curve in Fig. 1. This behavior arises from the spectrum (9)—the implication in the present treatment that there is a minimum excitation energy $2\hbar^2/md^2$ needed to affect the spin correlation, and hence to affect the susceptibility. In actuality, the energy spectrum of excitations of the liquid may be expected to extend continuously to the ground state; and one would expect to find some excitations which affect the spin correlations in any given energy range. Specifically, the present theory may be expected to fail at temperatures below the lowest reached in the recent resonance experiments because it leaves out of account long-range spin correlations. These will have to be dealt with by methods other than that used here. A noticeable feature of the foregoing results is that the theoretical $\chi(T)$ curve actually has a maximum. This occurs at $T/T_0=0.93$, and the maximum value is given by $(T_0/T)F\equiv(kT_0/N\mu^2)\chi=0.73$. The value of T_0 adopted places the theoretical maximum at 0.36°, within the temperature range of the observations. The experimental values of χ, however, show no such maximum. This disagreement merely illustrates how our theoretical $F(T)$ falls off too steeply at the lower temperatures, as discussed above.

It is of interest to apply the foregoing ideas to solid He^3. Pomeranchuk[4] has proposed that, as in an ordinary solid, there are no spin-ordering exchange forces and

hence (above about 10^{-7} degree, where the actual magnetic interaction becomes effective) the spins are uncorrelated and their distribution contributes R ln2 to the molar entropy, and χ is given by (1). The matter is not obvious, however, because the liquid and solid phases of helium resemble each other, in some ways, more than do those of other substances. In both phases the density is controlled principally by the zero-point kinetic energy. In the solid, although the density presumably varies in space with a crystal periodicity, yet, equally with the liquid, the amplitude of atomic motions is comparable with the interatomic distance: the atoms are continually colliding, rather than being kept away from each other by the conventional elastic forces.[8] We might therefore expect exchange forces to operate and cause spin correlations. It is doubtful, however, whether in the solid a neighboring pair of atoms may freely rotate completely about each other so as to change places; and it is just this free rotation of pairs which led to a separation of singlet and triplet spectra, and hence to spin correlations, in the foregoing theory for the liquid. We may expect rotation of neighbors in the solid to require a minimum excitation energy; so that at the lower temperatures it does not happen, and Pomeranchuk's conclusion is correct. One may speculate that appreciable rotation does occur at higher temperatures, near the rising part of the melting curve: if this is true then F should fall somewhat below unity, the spins being correlated by the same mechanism as for the liquid. We would then expect that, if the paramagnetic susceptibility is measured at temperatures increasing through the melting point, say at a constant pressure of 30 atmospheres, $F(T)$ would be found to dip slightly (by at most about 10 percent) from unity to a minimum somewhat below the melting point.[9]

I am obliged to P. Linhart for computing the numerical results used here, and to Miss K. Kaszas for drawing the figure.

APPENDIX

Equation (3) may be proved as follows. The energy levels of the whole system may be classified according to the eigenvalue, as M', of M_z. We may form separate partition sums for each value M': in zero field,

$$Z_{M'}(0)=\sum_{M_z=M'} \exp(-\beta E), \quad \beta\equiv 1/kT. \tag{A1}$$

The complete partition sum is

$$Z(0)=\sum_{M'} Z_{M'}(0). \tag{A2}$$

In a field H in the z direction, the energy E of a stationary state becomes $E-M'H$ (since we are ignoring the diamagnetism of the atomic electron shells). Then

$$Z_{M'}(H)=Z_{M'}(0)\exp(\beta M'H), \tag{A3}$$

$$Z(H)=\sum_{M'} Z_{M'}(0)\exp(\beta M'H). \tag{A4}$$

The mean value of M_z is then

$$\langle M_z\rangle=\sum_{M'} M' Z_{M'}(H)/Z(H),$$
$$=\sum_{M'}\beta H M'^2 Z_{M'}(0)/Z(0)+O(H^3)$$

in small fields. Hence

$$(d\langle M_z\rangle/dH)_{H=0}=\beta\langle M_z^2\rangle_{H=0}, \tag{A5}$$

which proves the result.

Note added in proof.—Since submitting this paper I learned that O. K. Rice has independently proposed essentially the same theory (see the following paper). His result (17) differs from my (8), (10), (11) in his use of only the lowest singlet level and lowest triplet level of the rigid rotator spectrum given by my (9). Professor Rice's paper differs from mine by his thorough analysis of possible systems of two-atom states, and stimulating discussion of the physical status and implications of these ideas. Rice's two-level formula (17) gives a significantly better fit than mine to the experimental data: one may remark about this that it is physically plausible that the pair model would break down for the highly excited states. If not, then in any case for these states the levels should not be ordered as in the scheme (9), which excludes excitations of radial motion. We both found that anomalously large values of d (Rice's γ_0) are needed to fit the data: about this one may remark (a) that for a state of nonzero angular momentum the cell should be flattened along the rotation axis, so that for fixed cell volume d is larger and the energy lower, and hence the level spacing (in particular between the first two levels) is reduced; and (b) that if there is some correlation between the motions of the pair and of the surrounding atoms then, for fixed density, the average pair separation should be increased. The difference in the potential energy expectation of the pair, and in the effect of pair motion on the energy of the surrounding atoms, between the levels might, however, be as important as these effects. In any case not too much should be expected of the pair approximation (6), I believe, and hence it soon ceases to be worth while trying to improve on models of the pair states, such as that leading to (10). There is, however, an interesting relevant development, which I learned of since submitting this paper: Temperley (private communication) has found that the experimental entropy and specific heat of the liquid[10] may be represented by a partition sum corresponding (apart from a factor corresponding to a Debye spectrum) to a system of independent pairs each with a spectrum consisting of a single ground level, a triple level up to 0.4°, and an octuple level up a further 2.3°. However, leaving aside the upper level, the fact that the middle level is only three-fold degenerate instead of nine-fold (as the first ortho-level; or five-fold as the second para-level) seems quite mystifying.

[8] This was established for He^4 by F. London (reference 6). His conclusion applies equally to He^3.

[9] It is possible that the relaxation time in the solid would be inconveniently long for the absorption method. Equilibrium being once established, the free precession signal, or a spin echo, due to a "90° pulse" could still be observed.

[10] B. Weinstock, Chicago A.P.S. meeting, November 1954.

Interpretation of the Magnetic Behavior of Liquid Helium-3*

O. K. RICE
Department of Chemistry, University of North Carolina, Chapel Hill, North Carolina
(Received October 6, 1954)

Recent work indicates that the lining-up antiparallel of nuclear spins in liquid He^3 takes place at temperatures much lower than that which would be expected for an ideal Fermi-Dirac gas. This indicates that the energy gap between the state in which the spins are aligned and one in which they can orient themselves in a magnetic field is lower than in the ideal gas. A model is considered in which pairs of atoms rotate about each other, and it is shown that the energy gap will be less for rigid spheres than for point particles, which more resemble the ideal-gas case. The model based on rigid spheres gives a reasonably shaped curve for the temperature dependence of the magnetic susceptibility, but the energy gap is still too large. If the rotation is hindered the gap is lowered, and a rough estimate of the degree of hinderance of rotation in liquid He^3 could be made. However, it appears likely that cooperative phenomena involving all the atoms play a role; these are discussed and a comparison with He^4 is made.

FAIRBANK, Ard, and Walters[1] have recently measured the nuclear magnetic susceptibility of liquid He^3 below 1°K. At very low temperatures the nuclear spins of the He^3 tend to line up antiparallel and this causes the spin magnetic susceptibility to deviate from the classical $1/T$ law. Such a behavior would be expected of an ideal Fermi gas; indeed in an ideal Fermi gas of the same density and atomic mass as He^3 it should occur at a temperature ten times as high as that at which it is actually observed in liquid He^3. In solid He^3, where each atom is more nearly in its own cell and there is certainly much less exchange of atoms or sharing of volumes, it might be that this lining up of the spins would not occur.[2] Thus liquid He^3 represents an intermediate condition, as might, indeed, be expected.

Some idea of the mechanism by which such an intermediate situation may arise can be obtained by considering the energy level of a pair of atoms, treating them first as more or less hard spheres and then as point particles. The case of "hard" spheres is of course more like the situation which would actually occur in liquid He^3, and we shall attempt to set up a model which can be used as the basis for a semiquantitative description of the actual behavior of He^3.

Let us consider the motion of a pair of He^3 atoms situated in He^3 liquid. As an approximation we shall suppose that the center of gravity of this pair of atoms is fixed at the center of a cell in the liquid, and that they can rotate and vibrate like a diatomic molecule. Suppose the effective diameter of the atoms to be σ and their average distance from each other to be r_0. The effective range of radial motion in the liquid will then be approximately $2(r_0-\sigma)$, and these "vibrations" may be supposed to carry them between the relative distances σ and $2r_0-\sigma$.

The symmetry properties of the system are dependent upon the rotational state. If the rotational quantum number, j, is zero, then the nuclear spins must be paired antiparallel, and such a state will not contribute to the magnetic susceptibility. If $j=1$, then the spins are in the same direction, and this state can contribute to the magnetic susceptibility. The rotational energy of this state is given approximately by the usual formula,

$$\epsilon_j=j(j+1)h^2/8\pi^2\mu r_0^2, \tag{1}$$

where μ is the reduced mass of a pair of atoms (equal to one-half the atomic mass m). For the energy of excitation necessary to get the system in a state which can contribute to the magnetic susceptibility, we obtain

$$\epsilon_1-\epsilon_0=h^2/4\pi^2\mu r_0^2=h^2/2\pi^2 m r_0^2. \tag{2}$$

We wish to compare this situation with that which occurs when the atoms are considered as point particles. In doing this, we need to describe a system of point particles which is as nearly comparable to the pair of hard spheres as is possible. We are interested in the interaction of the two particles with each other, not with the other particles, and we shall consider a pair of point particles which can range over the whole volume occupied by the equivalent hard spheres, namely a volume of diameter $2r_0$. The two particles have independent energy levels. If both particles are in the lowest level, the spins are paired, and no contribution is made to the magnetic susceptibility, but if one of them is excited such a contribution is possible. The energy levels of the atoms are best found from the wave equation in spherical form, using the polar coordinates, r, θ, and ϕ. Just as with the hydrogen atom, θ and ϕ equations can be separated, and the r equation takes the form

$$\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{dR}{dr}\right)-\frac{l(l+1)}{r^2}R+\frac{8\pi^2 m}{h^2}(E-V)R=0. \tag{3}$$

Here l is the rotational quantum number, R is the radial part of the wave function, and V, the potential

* Work supported by the Office of Naval Research.
[1] Fairbank, Ard, and Walters, Phys. Rev. 95, 566 (1954).
[2] I. Pomeranchuk, J. Exptl. Theoret. Phys. (U.S.S.R.) 20, 919 (1950).

energy, has the form

$$V=0 \quad \text{for } r<r_0$$
$$=\infty \quad \text{for } r>r_0. \tag{4}$$

Solutions of this equation for the region $r<r_0$ which have the proper behavior at $r=0$, are given[3] by the Bessel functions of order $l+\frac{1}{2}$. The first two solutions ($l=0$ and $l=1$) are

$$x^{-\frac{1}{2}}J_{\frac{1}{2}}(x)=(2/\pi)^{\frac{1}{2}}x^{-1}\sin x,$$
$$x^{-\frac{1}{2}}J_{\frac{3}{2}}(x)=(2/\pi)^{\frac{1}{2}}x^{-1}(x^{-1}\sin x-\cos x), \tag{5}$$

where

$$x=(8\pi^2mE/h^2)^{\frac{1}{2}}r. \tag{6}$$

Therefore the eigenvalues for $l=0$ are given by the positive roots of the equation

$$\sin x_0=0, \tag{7}$$

where x_0 is the value of x when $r=r_0$. The first of these is $x_0=\pi$, which gives

$$E_{01}=h^2/8mr_0^2. \tag{8}$$

The eigenvalues for $l=1$ are given by the positive roots of the equation

$$\tan x_0=x_0. \tag{9}$$

The first root of Eq. (9) is $x_0=4.493$ (see reference 2, p. 84). Thus the first eigenvalue for $l=1$ is

$$E_{11}=4.493^2h^2/8\pi^2mr_0^2. \tag{10}$$

This represents the lowest excited state. If one of the particles is in this state, the two particles can have the same spin and can contribute to the magnetic susceptibility. The excitation energy, therefore, is

$$E_{11}-E_{01}=[(4.493/\pi)^2-1]h^2/8mr_0^2. \tag{11}$$

Since $(4.493/\pi)^2-1\sim 1$, we see that

$$(E_{11}-E_{01})/(\epsilon_1-\epsilon_0)\sim\pi^2/4\sim 2.5. \tag{12}$$

Thus higher excitation for the point particles is required; therefore the magnetic susceptibility will fall below that anticipated from the Curie law at a higher temperature.

The difference between Eqs. (2) and (11) can be explained in a simple intuitive manner. In the case of the hard spheres, the small values of r are always excluded. This is true for both the lowest and next lowest energy level. In the case of the point particles, however, the small values of r are allowed for the lowest level, but for the next level, with $l=1$, they are effectively excluded by the rotational-potential term $l(l+1)/r^2$. This decrease in the free radial configuration space results in a relatively greater increase in the momentum, hence in the energy. It is important to point out that the effect of changing a pair of point particles to a pair of hard spheres is really to raise the lowest energy level more than the higher energy levels are raised. This may be a little difficult to see when one considers only the rotational energy, i.e., the energy connected with the angular momentum of the quantum state. It is evident, however, that the zero-point "vibrational" energy (i.e., that connected with the radial motion) is much higher for the hard spheres than for the point particles, and then it is seen that the smaller *difference* between the rotational states of the hard spheres merely means that the higher energy states are not raised as much above those of the point particles as is the case with the ground state.

We shall now calculate the magnetic susceptibility, using a model in which we suppose that N atoms can be considered as $N/2$ independent pairs of hard spheres. This, clearly, is not too realistic, especially because the pairs cannot be independent. We may hope, however, to make a calculation which can lay the groundwork for a semiquantitative understanding of the situation.

We shall consider only the two energy levels, with $j=0$ and $j=1$ (assuming always that we are dealing with the lowest vibrational state). Let us suppose that the system is in a magnetic field H, and let the component of the magnetic spin moment of an He^3 atom be either M, if in the direction of the field (energy $=-MH$), or $-M$ if in the opposite direction. If the system is in the lower energy level ($j=0$) the resultant moment of a pair is zero. In the higher triplet energy level the component of the magnetic moment of a pair is $2M$, 0, or $-2M$. The Boltzmann factors for these energy levels ($j=1$) are, respectively, $\exp(-h^2/2\pi^2mr_0^2kT+2MH/kT)$, $\exp(-h^2/2\pi^2mr_0^2kT)$, and $\exp(-h^2/2\pi^2mr_0^2kT-2MH/kT)$, and the quantum weight of each of these states is $2j+1=3$. The total magnetic moment induced in N atoms or $N/2$ pairs will then be

$$\chi H=\frac{3(N/2)(2M)[\exp(-h^2/2\pi^2mr_0^2kT)][\exp(2MH/kT)-\exp(-2MH/kT)]}{1+9\exp(-h^2/2\pi^2mr_0^2kT)}, \tag{13}$$

where χ is the susceptibility of the N atoms. The factors arising from the magnetic energies are small if H is small, and are omitted from the partition function [denominator of Eq. (13)]. Assuming $2MH/kT\ll 1$ and expanding the corresponding exponentials in Eq. (13), we obtain

$$\chi=\frac{(12NM^2/kT)\exp(-h^2/2\pi^2mr_0^2kT)}{1+9\exp(-h^2/2\pi^2mr_0^2kT)}. \tag{14}$$

At very high temperatures this reduces to a Curie law of the form

$$\chi_0=12NM^2/10kT. \tag{15}$$

[3] E. P. Adams, *Smithsonian Mathematical Formulae and Tables of Elliptic Functions* (Smithsonian Institution, Washington, D. C., 1922), pp. 202–203.

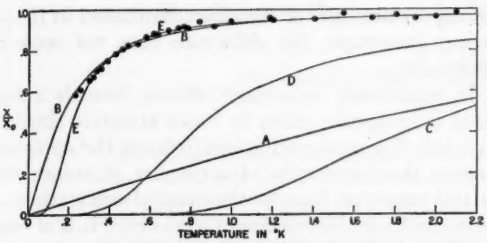


FIG. 1. Relative nuclear magnetic susceptibility as a function of temperature. Curve *A*, ideal Fermi-Dirac gas with density and atomic mass of He³. Curve *B*, ideal Fermi-Dirac gas with adjusted parameter. Curve *C*, pair model, point particles. Curve *D*, pair model, hard spheres. Curve *E*, pair model, hard spheres with adjusted parameter. The experimental points and curves *A* and *B* are from the figure of Fairbank, Ard, and Walters,[1] an enlarged copy of which was kindly furnished by Professor Fairbank.

On the other hand, if the atoms were completely independent the limiting high-temperature form would be

$$\chi_0 = NM^2/kT. \tag{16}$$

The difference between Eqs. (15) and (16) is due to the use of only two rotational levels to obtain (15). It seems, therefore, reasonable to set as a rough approximation:

$$\chi/\chi_0 = 10\exp(-h^2/2\pi^2 m r_0^2 kT) \times[1+9\exp(-h^2/2\pi^2 m r_0^2/kT]^{-1}. \tag{17}$$

We may first test the shape of this curve by giving mr_0^2 an arbitrary value so as to best fit the data. In Fig. 1, we show such a comparison with the experimental results, and also a curve for an ideal Fermi-Dirac gas with the constant similarly adjusted. It is seen that Eq. (17) gives a curve which is quite similar to that of the Fermi-Dirac gas except at very low temperatures, and that it fits the experimental data at least as well.

It is most interesting, however, to see what we obtain by setting m and r_0 equal to values expected from He³. When this is done for the ideal Fermi-Dirac gas (using the known value of m and of the density) and for Eq. (17) (setting r_0 equal to the corresponding distance in He⁴, 3.15A, known from x-ray data,[4] times the cube root of the ratio of the density[5] of He⁴ to that[6] of He³), it is seen (Fig. 1) that Eq. (17) with these parameters gives a curve which is better than that for the ideal Fermi-Dirac gas but which is still not very close to the experimental data.

A model using pairs of independent point particles, as in Eq. (11), gives a curve rather close to that for the ideal Fermi-Dirac gas. This is also shown in Fig. 1. Account has been taken of the difference in the possible spin states in this case, from those in the case of hard spheres, in drawing this curve.

[4] J. Reekie and T. S. Hutchison, Phys. Rev. **92**, 827 (1953).
[5] W. H. Keesom, *Helium* (Elsevier Publishing Company, Amsterdam, 1942), p. 207.
[6] E. C. Kerr, Phys. Rev. **96**, 551 (1954).

The fact that Eq. (17), with the proper parameters inserted, gives a curve which is still not very close to the experimental data, indicates that the energy gap between the states with symmetrical and antisymmetrical orbital wave functions, given by Eq. (2), is still too high. The free rotation, therefore, still represents a situation in which it is too easy for the atoms to change places. If we assume that the rotation were forced to take place in a plane rather than taking place freely in space, this halves the energy gap, because the formula for the energy of a plane rotation depends on the rotational quantum as j^2 rather than as $j(j+1)$. This still does not suffice to bring the curve into coincidence with the experimental points.

If the rotation were hindered this would lower the energy gap still further. For example, a potential barrier at rotational angles of 0 and π would cause the first two wave functions to have forms of the type shown in Fig. 2. The symmetrical wave function acquires what amounts to a superimposed wavelength of half its original wavelength, which means that its energy is raised, and the energy gap is thus lowered. It is, in fact, well-known that the energy gap is roughly inversely proportional to the time required for a rotation to occur, so one could use the results on the magnetic susceptibility to make an estimate of the extent of the freedom of mutual rotation of atoms in He³. The model, however, is probably not good enough for one to take such an estimate too seriously.

As a matter of fact, comparison with He⁴ indicates that the situation must be more complicated than would be inferred from this simple picture. Direct calculation from Eq. (2) indicates that the energy gap, with the proper parameters for He³ inserted, is about 4.5 calories per mole, or about 2.2 calories per mole of He³ atoms. In order to fit the data on the magnetic susceptibility, the energy gap should be about $\frac{1}{4}$ of this, or 0.5 calorie per mole. This is a very small excitation energy; in He⁴ the roton excitations involve approximately 2.5 calories per mole[7] of excited He⁴ atoms. This difference

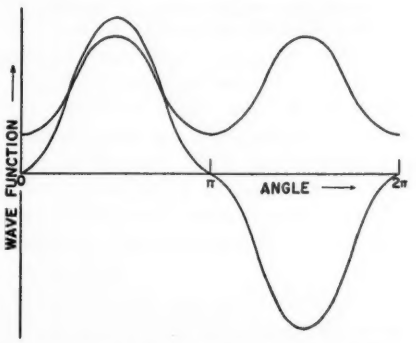


FIG. 2. The two lowest rotational wave functions with hindered rotation.

[7] See O. K. Rice, Phys. Rev. **96**, 1460 (1954).

may be, at least in part, connected with the fact that the quantum number j cannot have the value 1 in the rotation of a pair of He^4 atoms, on account of the statistics. Insofar as the pair approximation is good, then, the first excitation in He^4 must be to the next higher level, $j=2$. However, it is also probable that a cooperative interaction of all the atoms, peculiar to the Fermi-Dirac statistics, has something to do with the very low excitation energy in the case of He^3.

Let us consider the situation at 0°K, where presumably one-half the atoms have their spins in one direction and half have their spins in the opposite direction. Since the substance is in the liquid state and since the zero-point motion is considerable, it is hardly possible that the atoms have reached a state at 0°K in which no exchanges occur. Indeed, if they had, the spin-antiparallel alignment would have been destroyed. But if exchanges occur, we must assume that they take place between atoms which have mutually symmetrical spin functions (symmetrical exchanges) as well as between atoms which have antisymmetrical spin functions (antisymmetrical exchanges), and there must be some of each since half of the atoms have the same spin. However, if symmetrical exchanges were as frequent as antisymmetrical exchanges, there would be a considerable amount of localized angular momentum, which would force up the energy of the whole system. In order to avoid this situation, there will be some tendency for atoms in symmetrical spin states to stay as far away as possible from each other. This, however, raises the zero-point energy, so some compromise will have to be found. In any event, it is seen that the energy of the whole system is higher than it would be if the exchange problem were not involved,[8] (i.e., if there were no spin, and all exchanges were symmetrical as in He^4). Since the ground level of He^3 is thus raised, this may be another reason why excitation to a higher state should be lower in He^3 than in He^4. It may seem that the difference between 0.5 cal per mole and 2.5 cal per mole of excited atoms is large, but these energies are very small compared to the zero-point energy, so, taking account also of the effect mentioned in the preceding paragraph, the difference does not seem unreasonable.

In considering cooperative effects, there is a somewhat more specific point to which attention should be directed. The roton excitations in liquid He^4 apparently require the cooperation of a number of atoms, while in this paper, at least in the detailed calculations, we have considered interaction of pairs only.[9] It is of course possible that excitation in He^3 requires cooperation of several atoms, also. However, in the light of the lower density of He^3, and the presumed preliminary "excitation" even at absolute zero, described above, it may be that the consideration of pairs only is a better first approximation for finding the energy of excitation in this case than it would be in the case of He^4.

It is clear that the model used here can be made the basis of a theory of the specific heat of liquid He^3.

Note added in proof.—After submission of this paper I learned that P. J. Price[10] had already submitted a paper in which the same problem was treated by the same independent-pair model. The two papers appear to supplement each other; there is some difference in emphasis, for Price has stressed the possibility of finding a more reasonable model than the ideal-Fermi-gas model to account for the magnetic results, whereas I have particularly endeavored to find a semiquantitative explanation for the difference between the observed temperature parameter and that predicted from the ideal gas. Price has worked out the consequences of the model using the full set of energy levels; the fact that I find in better agreement with the experimental results by considering only the two lowest energy states may be a reflection of the hindered character of the rotation of a pair, the effect of the surrounding molecules being to bring the lowest levels closer together and effectively to isolate them from the others. Price mentions the possibility that long-range interactions will cause deviations from the calculated susceptibilities at very low temperatures; if these deviations are *not* too important, He^3 may possibly exhibit superfluidity at sufficiently low temperatures.[11]

[8] In this connection, compare my earlier remarks, O. K. Rice, Phys. Rev. 93, 1161 (1954), Sec. 4.

[9] I. Prigogine, J. Philippot, and co-workers [see I. Prigogine, Advances in Physics 3, 131 (1954)] have obtained qualitative agreement with the properties of liquid He^4 considering only the interaction of pairs of atoms. However, there seems to be little doubt that the elementary excitations actually involve several atoms. Their results for a Fermi gas apparently do not coincide with those found here.

[10] P. J. Price, Phys. Rev. 97, 259 (1955).

[11] O. K. Rice, Phys. Rev. 97, 263 (1955).

Breakdown Processes in Nitrogen, Oxygen, and Mixtures*†

ELSA L. HUBER
Department of Physics, University of California, Berkeley, California ‡

(Received October 13, 1954)

The studies of corona mechanisms in coaxial cylindrical geometry with central anode using α-particle triggering, initiated by Colli and Facchini and by Lauer in inert gases, have been extended to pure nitrogen, pure oxygen, and mixtures of these gases. Pure gases were studied in a well outgassed system with nickel cathode. In pure nitrogen, from 25 to 600 mm pressure, the predominant mechanism is secondary liberation of electrons from the cathode by positive ion impact, with γ_i varying between 1.4×10^{-3} and 0.6×10^{-3}. Correction for back diffusion following Theobald gives a true γ_i of $2\text{–}3\times10^{-3}$. The ion is the N_4^+ observed by Varney, with reduced mobility 2.53 ± 0.08 cm²/volt sec at 760 mm and 20°C. Addition of the order of 1 percent oxygen reduces γ_i by a factor of $\frac{1}{3}$ and gives a weak photoelectric γ_p at the cathode, both effects resulting in part from action of oxygen on the cathode as noted by Parker and Theobald. Photoionization *in the gas* also appears, and with 5 percent oxygen the self-sustaining corona consists to a large extent of Geiger counter-like pulses propagating along the wire. In air, there is no longer any observable cathode action at threshold. With the use of transverse and longitudinal α particle trajectories relative to the axis, the velocity of propagation of the discharge down the wire has been measured for various pressures in air, and for constant pressure and potential as a function of oxygen concentration of between 5 and 20 percent. Velocities ranged between 10^8 and 10^7 cm/sec, decreasing as oxygen concentration increased. In pure oxygen above 200-mm pressure, the absorption of photoionizing radiation is so intense that, as indicated by Miller and Loeb, instead of spreading along the wire, the discharge propagates outward in the form of radial streamers, which have a duration of less than 10^{-7} sec, contain some 10^8 ions, and are spaced at regular intervals closely related to the positive-ion transit time. The latter yields a reduced mobility of 2.2 ± 0.1 cm²/volt-sec at 760 mm Hg and 20°C. in conformity with Varney's assumed O_2^+ ion.

INTRODUCTION

THE method developed by Lauer[1] from an earlier study of Colli and Facchini[2] for investigating the discharge mechanisms in a coaxial cylindrical system, with central wire positive, involved triggering the discharge by single α particles projected parallel to the axis of the electrodes. Such dynamic studies have proved so fruitful because of their high resolving power that the technique was extended to nitrogen, oxygen, and mixtures thereof. As had previously been shown with more primitive technique by Miller and Loeb,[3] a wide variety of secondary mechanisms appear to occur in these gases. The excellent resolving power of the Lauer techniques thus offered a unique opportunity for direct visual and even quantitative study of the amazing sequence of breakdown phenomena in O_2 and N_2 mixtures as O_2 concentration increases.

APPARATUS AND PROCEDURES

The discharge tube was essentially that used by Lauer with the modifications noted. In order to minimize any leakage currents between the positive wire and the grounded cylinders via poor insulators, the platinum anode wire was supported only from the ends of the glass envelope, with a tungsten spring at the lower end to maintain tension on the wire. After the cylinders had been mounted in the tube, the wire was aligned visually with two sets of mutually perpendicular marks on the outer rims of the two guard cylinders. A tungsten filament was provided outside the cathode cylinders for outgassing them by heating by electron bombardment.

The discharge was triggered by either of two α-particle sources. One source (longitudinal) furnished a beam of α particles parallel to the axis of the system and passing through the low-field region near the cathode. A second source (transverse) directed a beam between the upper two cylinders, perpendicular to the system axis but likewise passing through the low field region only. The α sources were constructed by plating polonium on the tip of a silver wire and inserting the wire 1 cm from the open end of a capillary tube 0.57 mm in diameter. This arrangement provided good collimation for the short trajectory of the transverse beam. For the longitudinal beam, a maximum deviation from the vertical of 3 mm was possible and could in some cases give rise to a slight variation in shape from one pulse to the next. Each source emitted approximately one α particle per second on the average. Pyrex shutters to which were attached slugs of soft iron could be operated by a small permanent magnet to turn the beams on and off.

For work with the pure gases nitrogen and oxygen, the discharge tube, which was mounted in an oven, was baked out at 450°C at least overnight, with mercury diffusion pump in continuous operation. Outgassing of the nickel cylinders consisted of heating them in

* This work was supported at various stages by both the U. S. Office of Naval Research and the National Science Foundation.

† This work has been presented in more detail in the Ph.D. thesis of Elsa L. Huber, Department of Physics, the University of California, Berkeley, California.

‡ Now at The University of California Radiation Laboratory, Livermore, California.

[1] E. J. Lauer, J. Appl. Phys. **23**, 300 (1952).

[2] L. Colli and U. Facchini, Phys. Rev. **88**, 987 (1952); L. Colli, Phys. Rev. **95**, 892 (1954); L. Colli and U. Facchini, Phys. Rev. **96**, 1 (1954).

[3] C. G. Miller and L. B. Loeb, J. Appl. Phys. **22**, 494 (1951); **22**, 614 (1951); **22**, 740 (1951).

vacuum to a bright red heat for many hours by electron bombardment from the tungsten filament. The platinum filament was heated to a dull red for a few minutes by passing a current through it. The system was flushed with the purified gas to be used before filling.

Nitrogen was admitted into the system via a slow leak, then purified by passing it over potassium hydroxide pellets, then hot copper shot, and finally through a liquid nitrogen trap to remove carbon dioxide, oxygen, and water vapor, respectively. No difference in behavior was noted between Airco Prepurified Nitrogen (less than 0.002 percent oxygen) and Ohio's ordinary commercial variety when both were subjected to the purification described. Analysis by mass spectrometer of two samples of gas taken from the tube after purification indicated that flashing the W filament as a getter introduced more impurity than it removed.

In the case of oxygen extreme purity was less critically important than for nitrogen, so that after bakeout, outgassing, and flushing, ordinary tank oxygen was admitted at a fairly rapid rate, with the copper shot cold. Liquid nitrogen on the cold traps was replaced by dry ice-acetone baths.

The system was filled with room air also without any special precautions, except that again a dry ice-acetone mixture was used on the traps instead of liquid nitrogen to avoid condensing any of the oxygen. In addition to keeping mercury vapor out of the discharge tube, the cold traps helped insure uniform water vapor content of the air from one day to the next.

The signal from the discharge tube was picked up across a variable RC from the central collector cylinder to ground, amplified and observed on an oscilloscope, then recorded photographically. A Tektronix Type 511A synchroscope and Type 121 preamplifier (6×10^{-8} sec rise time for the combination, 1×10^{-7} sec/cm maximum sweep speed) were adequate for most of the work. For the spreading velocity studies in air, which required a more rapid response, it was necessary to use the Type 517 Tektronix synchroscope, in conjunction with two Model 460A Hewlett Packard wide band amplifiers. The net rise time with this arrangement was 8×10^{-9} sec, with a maximum possible sweep speed of 10×10^{-9} sec/cm.

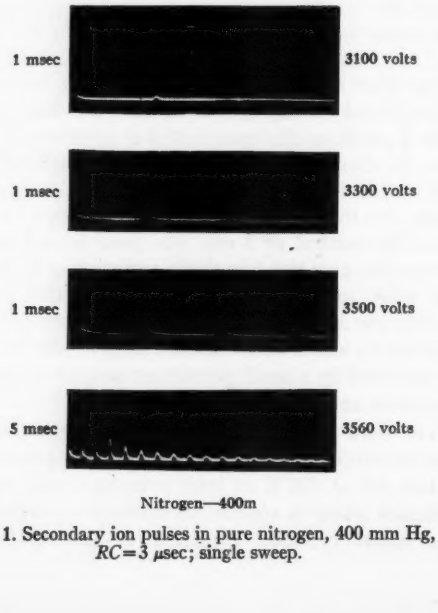


FIG. 1. Secondary ion pulses in pure nitrogen, 400 mm Hg, $RC=3$ μsec; single sweep.

CHARACTERISTICS OF SECONDARY MECHANISMS

Reference may be made to Lauer's paper for details of some of the sequence of mechanisms observed. To facilitate the reading certain features must be recalled. The α particle creates n_0 electrons and ions near the outer cathode cylinder. At appropriate fields the fine linear electron cloud crosses to the anode in a transit time τ_e. On its way each electron produces an avalanche of $M=\exp(\int_b^a \alpha dr)$ electrons, where α is the first Townsend coefficient, r is the radial distance, and a and b the radii of anode and cathode respectively. Absorption of electrons by the anode discloses in the gap, but near the anode wire, the space charge of positive ions left by the avalanche. This induces a current flow from ground to cathode leading to an initial sharp pulse on the oscilloscope. This is the primary and triggering mechanism. The positive ions drift to the cathode in a time τ_i leading to a decline of the current pulse to zero on arrival. The avalanche of electrons, ions, and excited molecules results in three possible secondary mechanisms: (1) Positive ions on arrival at the cathode liberate electrons in virtue of their potential energy with a chance γ_i per ion. (2) Photons of sufficient energy which reach the cathode without resonance absorption liberate photoelectrons with a chance γ_p for the photon production accompanying each ion created in the avalanche. (3) Photons of sufficient energy but highly absorbed may cause photoionization of the gas which under appropriate circumstances will either propagate a discharge along the anode wire (*burst pulse*) or radially outward (*streamer*). Both of these self-sustaining discharges yield many electrons but ultimately quench by space charge. Process 1 is characterized by a succession of one or more secondary pulses following the primary one spaced $\tau_e+\tau_i$, or about a few hundred microseconds, apart. Process 2 is characterized by one or more secondary pulses spaced τ_e, or about a microsecond, apart. Process 3 will lead to pulses which are large, independent of triggering, above 100 microseconds long, and are irregularly spaced for burst pulses, but may be regularly spaced, very short, and with heavy ionization for streamers. Secondary liberation by entrapped radiation or by diffusion of metastable configurations to the cathode will be too slow and diffuse to record on the oscilloscope as used. Attachment of electrons to molecules to give negative ions will lead to irregular delayed pulses between τ_e and τ_i in time.

RESULTS AND DISCUSSION

Pure Nitrogen

The behavior of pure nitrogen was uniform over the entire pressure range studied, from 25 to 600 mm. As the voltage across the tube was gradually raised, the initial pulse representing the primary avalanche formation was followed by first one, then more and more secondary pulses. Figure 1 shows a representative set of oscillograms taken at a pressure of 400 mm. The secondary pulses were spaced at uniform intervals from the first at a given voltage and pressure. Consecutive pulses within one sequence showed a gradual decrease in amplitude, the decrease, however, becoming less pronounced as the onset voltage for a self-sustaining discharge was approached. The lower trace photograph was taken at 10 volts below onset.

Spacing between regular pulses corresponds to one ion transit time from wire to cylinder, indicating that the important secondary mechanism is electron liberation from the cathode by positive ion impact. With a faster sweep speed adjusted to τ_e, there was no indication of any photoelectric action at the cathode. The smooth shape of the ion pulses indicates secondary electron production in the gas to be negligible.

The spacing between pulses yielded the ion mobility. Data were reduced to mobilities as by Lauer, no space charge correction being required. The plot of the reciprocal mobility against pressure was strictly linear. Extrapolation of the resulting straight line gave a value of 2.53 ± 0.08 cm^2/volt-sec at 760 mm Hg and 20°C. The work of Varney on nitrogen mobilities[4] and the mass spectrographic studies of Luhr[5] indicate that the ion involved is N_4^+, a conclusion supported by the fit of the measured mobility to the mass dispersion curve for positive ions in nitrogen.[6]

From the oscillograms it was also possible to determine the magnitude of γ_i, the number of secondary electrons per incident positive ion released at the cathode as indicated by Lauer using n_0 and M.

The amplification factor $M=n/n_0$ was determined as follows: n could be obtained directly from the oscillograms by measuring the area under the primary pulse. At the maximum available gain, an observable deflection could not be obtained for the n_0 ions collected with no multiplication, and n_0 was therefore estimated from the known ionization of an α particle. Relative stopping power for nitrogen and oxygen are nearly the same (1.07 for oxygen with respect to 1.00 for air).[7] On the assumption that specific ionization is proportional to pressure, range-energy curves for α particles in air[8] were therefore used to find the energy lost over the 6-cm length of the center cylinder. This implies that where the track range extended beyond the center cylinder adequate correction was made. n_0 was then given by the ratio of this energy loss to the 35 ev required per ion pair.[7] Column (a) of Table I lists the resulting values of γ_i. Various other methods for calculating M, as indicated by Lauer and later by Colli and Facchini,[2] were used as checks. Accuracy at best was ±20 percent.

TABLE I. Secondary emission coefficient for nitrogen ions on nickel.

p(mm Hg)	(a) $1000\gamma_i$ (M_{meas})	(b) $1000\gamma_i$ (M_{calc})	(c) $(E/p)_{max}$	(d) $100j/j_0$	(e) $100\gamma_i'$
600	0.6	0.7	220	3.0	2.0
500	0.7	1.0	230	3.1	2.3
400	1.1	1.2	250	3.4	3.2
300	1.0	1.3	290	3.7	2.7
200	1.2	1.5	360	4.1	2.9
100	1.5	2.4	530	5.9	2.5
50	1.4	3.4	780	6.9	2.0

In this study, only electrons that escape back-diffusion to the cathode can participate in secondary avalanches and contribute to the measured γ_i. In order to determine the γ_i' that would be measured if all the secondary electrons escaped, correction was made following the procedure of Theobald.[9]

Column (d) of Table I lists j/j_0 as calculated from the Thomson[9] equation, and column (e) the corrected values of γ_i'. The data of Nielsen[10] were used to obtain the electron drift velocities v corresponding to the values of E/p at the cathode. The average energy of emission of the secondary electrons has been estimated at 3 ev, based on the mechanisms which have been proposed to describe electron ejection from a metal surface by slow positive ions.[11]

The corrected values of γ_i' could well be in error by a factor of 2 because of the approximations involved. No other measurements have been reported specifically for N_4^+ with thermal energy on nickel, but there is satisfactory agreement in order of magnitude with the results of other authors[12,13] for ions having ionization energies close to that of N_4^+, incident on metals with work functions similar to that of nickel.

Effect of Few Percent Oxygen in Nitrogen

With the addition of a few percent oxygen to the nitrogen, a number of distinct changes occurred:

1. The secondary pulses from γ_i were smaller and fewer in number at a given V and p than in pure nitrogen.

[4] R. N. Varney, Phys. Rev. 89, 708 (1953).
[5] O. Luhr, Phys. Rev. 44, 459 (1953).
[6] A. M. Tyndall, *The Mobility of Positive Ions in Gases* (Cambridge University Press, London, 1938).
[7] F. Rasetti, *Elements of Nuclear Physics* (Prentice-Hall Publications, New York, 1936).
[8] H. A. Bethe, "The Properties of Atomic Nuclei," AEC, Oak Ridge Tenn. (1949) (unpublished).
[9] J. K. Theobald, J. Appl. Phys. 24, 123 (1953).
[10] R. A. Nielsen, Phys. Rev. 50, 950 (1936).
[11] H. D. Hagstrum, Phys. Rev. 89, 244 (1953). Also private communication on need for heating to 900°C or more to remove contamination by exposure.
[12] J. H. Parker, Phys. Rev. 93, 1148 (1954).
[13] J. P. Molnar, Phys. Rev. 83, 940 (1951).

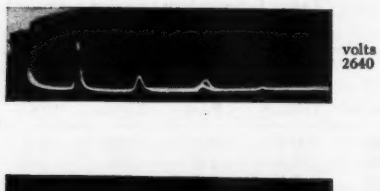


FIG. 2. Secondary ion pulses, top trace, and self-sustaining discharge in pure nitrogen, 200 mm Hg, 1000 μsec full scale, $RC = 3$ μsec. Upper trace—a few superposed sweeps. Lower trace—several single sweeps spearated by vertical displacement showing initial and secondary with self-sustaining fluctuation.

2. The secondary pulses from γ_i began to lose their smooth shape and became rather ragged.

3. Onset of a self-sustaining discharge occurred at a voltage approximately 10 percent higher than in pure nitrogen.

4. With a faster sweep speed and shorter circuit RC, a new type of secondary pulse was observed, spaced at intervals of about a microsecond from the primary.

The above effects are illustrated in the photographs of Figs. 2, 3, and 4, all taken at a pressure of 200 mm. Figure 2 shows the ion secondaries in pure nitrogen at

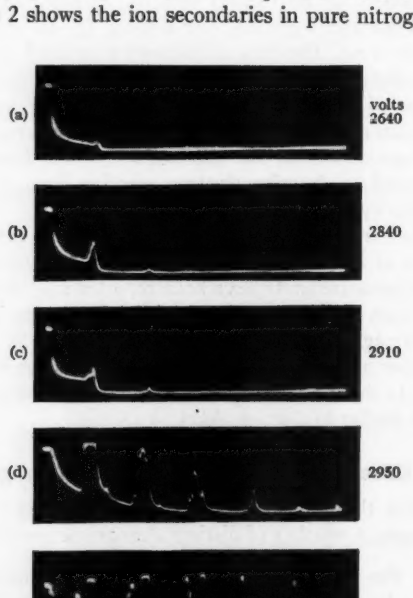


FIG. 3. Secondary ion pulses and self-sustaining discharge in nitrogen plus 5 percent oxygen, 200 mm Hg, 1000 μsec full scale, $RC = 3$ μsec. Traces such as d and e are single sweeps.

2640 volts and the self-sustaining discharge starting at 2660 volts. Comparison may be made with Fig. 3 for nitrogen plus 5 percent oxygen. In 3(a), barely one ion secondary can be seen at 2640 volts as opposed to four in pure nitrogen. Onset, as shown in 3(e), does not occur until V is raised to 2960 volts, 300 volts higher than in pure nitrogen. 3(b), (c), and (d) show how the ion pulses, while still clearly present, become progressively more irregular as the voltage is raised to onset. Compare the fluctuations of 3(e) with those in the lower trace of 2. The microsecond pulses are illustrated in Fig. 4.

It thus appears that there is a reduction in γ_i by a factor of approximately 5 with 5 percent O_2 over that of N_2. The change in γ_i is associated with an increase in the work function of the surface and a change in the secondary emission mechanism, as discussed by Parker,[12] who found at low ion energies for A^+ on platinum and tantalum that γ_i decreased by an order of magnitude if the cathode was treated with oxygen instead of nitrogen. That this change is a cathode phenomenon and not one due to possible change in ions was checked by removing the O_2 and refilling with pure N_2. Restoration of the initial γ_i on clean Ni did not occur until the cathode was heated to 900°C for hours in conformity to the experience of Hagstrum,[11] Parker[12] and recently Colli and Facchini[2] on this apparatus.

The microsecond pulses show the presence of a photoelectric γ_p in this mixture, absent in pure N_2, which is confirmed by the fact that the spacing of these fast pulses corresponds to an electron transit time in order of magnitude. These pulses are smeared and irregular compared to those of Lauer in H_2. It may therefore be concluded that these pulses represent secondary electrons released from the cathode by photons created in the preceding set of avalanches with, perhaps, some delay. Electron transit times were calculated by numerical integration of Nielsen's data[10] for electron drift velocity as a function of E/p in nitrogen.[14] Comparison between the measured τ_e, based on spacing between the microsecond pulses, and the calculated values is made in Table II. It is to be noted that these pulses, unlike those of Lauer in H_2, do not give a self-sustaining discharge at 2960 volts where the slower sweep pulses are self-propagating.

Absence of a cathode photoelectric effect in pure nitrogen may involve capture of photons by resonance absorption, or simply a difference in the photosensitivity of the cathode. The present results in any case agree with the observations of Burch, Irick, and Geballe, [15] who found an order of magnitude increase in the photoelectric yield from the cathode in a nitrogen-4 percent oxygen mixture and in air over that in either pure nitrogen or pure oxygen.

[14] To justify comparison with data taken in pure nitrogen, measurements were made on photoelectric pulses obtained with minimum possible oxygen concentration of about 1 percent.

[15] Burch, Irick, and Geballe, 1953 Conference on Gaseous Electronics, Washington, D. C., October 22–24, 1953 (unpublished).

The attachment of electrons to O_2 molecules to form negative ions and their subsequent detachment in the high fields near the anode to start delayed avalanches could account for an increase in pulse duration and small fluctuations in pulse amplitude. (See work of Lauer[1] in hydrogen with 0.1 and 1 percent oxygen.) However, the pronounced irregularities in pulse shape evident in Fig. 3(d), for example, cannot be explained in terms of negative-ion formation, but represent creation of secondary electrons by photoionization of gas molecules. With 5 percent oxygen, moreover, the oscilloscope pattern characterizing onset differs markedly from that in pure nitrogen. In Fig. 3(e) there is still some slight evidence of secondary ion pulses, but in addition, an indication that the discharge is being quenched, apparently along the entire length of the wire, at intervals shorter than an ion-crossing time. Such behavior would be characteristic of the burst pulse phenomenon described in detail by Loeb,[16] in which the discharge spreads along the wire by photoionization of gas molecules, burns until sufficient positive ion space charge accumulates to lower the field at the wire and decrease the multiplication, and thus is finally choked off. This action is also the one responsible for the discharge in fast Geiger counters. With clearing of the space charge the field recovers, and a new discharge can start somewhere along the wire length and again propagate along the high-field region by photoionization.

TABLE II. Comparison between measured and calculated electron transit times in oxygen-contaminated nitrogen.

p(mm Hg)	τ_e(μsec) Measured	Calculated
600	0.96	0.98
400	0.88	0.87
200	0.68	0.69
100	0.48	0.51
50	0.35	0.38

Nitrogen Plus 20 Percent Oxygen

Figure 5 illustrates another marked change in the character of the discharge mechanism that occurred when the oxygen concentration was increased to 20 percent. Below onset, the α-particle triggering produced irregular "bursts" of ionization, as shown in 5(a) and (b), which resemble the burst pulses first observed by Trichel[17] and Kip[18] in positive point-to-plane corona in air, and also noted by Miller[3] with cylindrical geometry. At onset the discharge went over into steady burst pulse corona, illustrated in Fig. 5(c).

There was no indication whatever of secondary electron production at the cathode, either by ions or photons. Nevertheless, the discharge became self-sustaining at

[16] L. B. Loeb, Phys. Rev. **73**, 798 (1948).
[17] G. W. Trichel, Phys. Rev. **55**, 382 (1939).
[18] A. F. Kip, Phys. Rev. **54**, 139 (1938); **55**, 549 (1939).

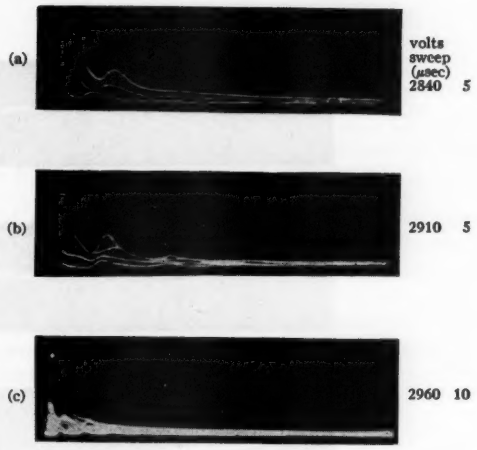


FIG. 4. Secondary photoelectric pulses in nitrogen plus 5 percent oxygen, 200 mm Hg, $RC=0.03$ μsec. Several sweeps indicate lack of accurate reproducibility and that at self-sustaining corona in trace (c) the photon γ_p does not sustain itself as a discharge.

2145 volts, appreciably lower than in either pure nitrogen or nitrogen with a few percent only of oxygen. This lowering of the onset potential indicates a marked increase in the production of secondary electrons in spite of the absence of a γ action at the cathode, the predominant secondary mechanism in this case being photoionization of gas molecules, with alternate propagation along the wire and space-charge quenching of the discharge, as described in the preceding section.

Spread of the Discharge Along the Wire in Air

The object of this part of the work was to establish that the spreading process does occur in dried room air,

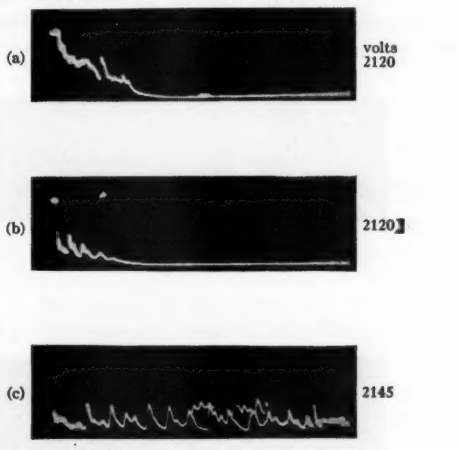


FIG. 5. Burst pulses and self-sustaining discharge in nitrogen—20 percent oxygen mixture, 200 mm Hg, 1000 μsec full scale, $RC=3$ μsec. Gain is down by a factor of 10 in (c). Single sweeps, except part of a second sweep in (c).

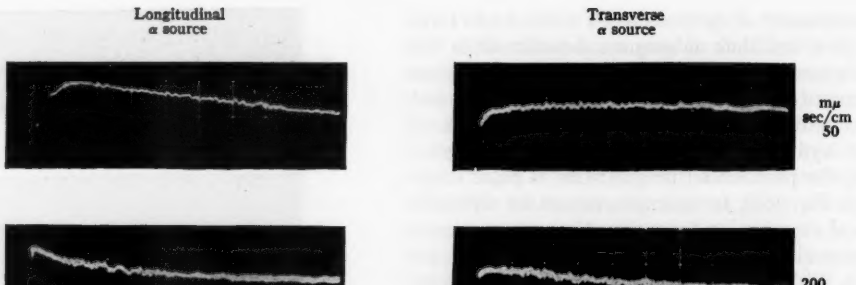


FIG. 6. Longitudinal and transverse α pulses in room air, 300 mm Hg, 2720 volts. Single sweeps at rates shown in margin

and to determine its velocity of propagation. Both the longitudinal α source (beam parallel to system axis) and the transverse source (beam perpendicular to system axis) were used for triggering purposes. In the case of the longitudinal source, the rise time of the pulse observed on the oscilloscope consists of the time for an α particle to travel the length of the cylinder, plus the time of formation of the electron avalanche created when the primary electrons ionize by collision in the high-field region near the central wire. With the transverse source, the pulse rise likewise contains the α-particle transit time, the time of formation of the initial avalanche, and in addition the time for the discharge to propagate down the length of the cylinder. The difference in rise time of the two types of pulses therefore yields the spreading time and, combined with the cylinder length, also the spreading velocity.

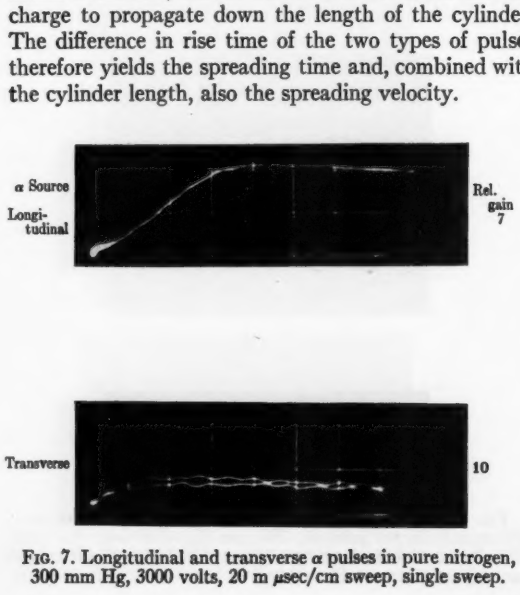


FIG. 7. Longitudinal and transverse α pulses in pure nitrogen, 300 mm Hg, 3000 volts, 20 m μsec/cm sweep, single sweep.

Figure 6 shows some typical pulses observed in air at a pressure of 300 mm and applied potential of 2720 volts. It can be seen that the rise time for the transverse pulse is indeed noticeably greater than for the longitudinal pulse. Furthermore, there is approximately the same amount of total charge (proportional to the area under the pulse) in the two types of pulse, in spite of the fact that the initial ionization produced by a longitudinal α particle is approximately an order of magnitude greater than that from a transverse α particle. This lack of proportionality between initial ionization and total charge produced is to be expected, because additional multiplication occurs in the case of the transverse pulse when the discharge spreads along the wire.

As further confirmation that the effect was real, a similar set of measurements was made in pure nitrogen, for which it has been established that ionization *in* the gas is negligible and in which propagation along the wire is therefore not possible. One would thus expect the rise times for the transverse and longitudinal pulses to be about the same, and this was observed to be the case, as illustrated in Fig. 7. Rise time of the longitudinal pulse in the photograph was, in fact, slightly *longer* than that of the transverse, presumably because of imperfect collimation.[19] Furthermore, without the additional ionization resulting from the spread of the discharge, the transverse pulse would be expected to contain a much smaller amount of total charge than the longitudinal. The total amount of charge contained in the two types of pulses was obtained by integrating the full pulses with slower sweep and was approximately proportional to the initial ionization in each case.

[19] Lack of perfect collimation of the longitudinal beam caused some scatter in pulse rise times. These rarely exceeded a few times 10^{-8} sec, however, and for purposes of calculation the fastest rise was taken as representative of best collimation.

TABLE III. Velocity of spread in air.

p(mm Hg)	V(volts)	v(cm/sec)
100	1580	3×10^7
200	2190	3×10^7
300	2720	4×10^7
400	3240	6×10^7

Table III gives the values of the velocity of spread measured in air at four different pressures covering the observation range. The measured velocities of from 10^7 to 10^8 cm/sec agree in order of magnitude with values obtained by other investigators,[20–22] usually in typical Geiger counter fillings.

The pressure variation of the velocity is rather difficult to interpret, especially since at different pressures it was necessary to operate at different potentials. The velocity of spread depends on the time of formation of an electron avalanche and therefore on both the electron drift velocity and the distance from the wire at which electrons can first begin to ionize by collision; on the average lifetimes of the molecules in excited states; and on the average length, i.e., the absorbing free path for ionizing photons, and thus the total number of individual fictive[23] ionizing steps in the process.[20] These factors are no doubt all functions of pressure and/or voltage and are therefore hard to separate.

Since oxygen is responsible for the ionizing absorption of photons, the air was diluted with either nitrogen or oxygen to give different concentrations of oxygen. Then at a fixed total pressure it was also possible to work at the same voltage. The only factor one would expect to change radically in this case is the mean free path for photon absorption, which should decrease with increasing oxygen concentration. The difference between the fixed potential and the threshold for self-sustaining discharge, which changes slightly with O_2 concentration, is a minor factor.[23] The net result of increased O_2 concentration is shorter ionizing steps, a greater number of total steps, and hence a slower velocity of spread.[23] Table IV shows the variation of the spreading velocity with percentage of oxygen. In pure nitrogen there is no spread, as has been pointed out; the velocity decreases steadily as the percentage oxygen increases from 5 percent to 10 percent to 21 percent; finally at 40 percent oxygen it was no longer possible to observe the α pulses before the discharge became self-sustaining, and apparently in this region there is a transition to the streamer mechanism typical of pure oxygen as will be indicated.

TABLE IV. Variation of spreading velocity with percentage O_2. $V=2720$ volts, $p=300$ mm Hg.

Percent O_2	v(cm/sec)
0	No spread
5	9×10^7
10	6×10^7
21	4×10^7
40	No α pulses

FIG. 8. Ion drift pulse in oxygen with amplification but with no secondary action, 300 mm Hg, 2775 volts, 1000 μsec full scale, $RC=3$ μsec.

Pure Oxygen

The nature of the various phenomena in pure O_2 seen on the oscilloscope as the voltage was raised is illustrated in Figs. 8–11 inclusive, for which the pressure was 300 mm.

Figure 8 shows the type of pulse observed when the applied voltage was relatively low—i.e., Townsend avalanches relatively small and little secondary mechanism. The duration of these pulses, as in N_2, yields the mobility of the ions in O_2, although the measurement is somewhat less precise than in nitrogen. The mobility extrapolated to 760 mm. and 20°C was 2.2 ± 0.1 cm^2/volt-sec. This is in satisfactory agreement with Varney's zero-field mobility of 2.25 ± 0.1 (at 0°C) as extrapolated from his data at higher E/p, for an ion he believes to be O_2^+ in accordance with Luhr's[5] mass spectrographic studies.

A secondary mechanism became apparent, as in Fig. 9, at a somewhat higher voltage. These pulses again are of the *burst* pulse variety, and were also observed by Miller in pure oxygen at pressures of 200 mm and less. They differ from those of Fig. 8 in that their amplitude is higher, shape more irregular, and duration somewhat longer. The fluctuations are of much shorter duration and are orders of magnitude smaller than the burst pulses of Fig. 5. Electron attachment to oxygen molecules near the cathode and subsequent detachment in the high-field region might account for the extension in duration and for some irregularity in pulse shape. The observed irregularities, however, are in principle localized discharges of the burst pulse type, triggered by negative ions but unable to propagate down any length of wire because of high absorption of photons. The large numbers of negative ions could not by

FIG. 9. Burst pulses in oxygen, 300 mm Hg, 2900 volts, 1000 μsec full scale, $RC=3$ μ sec. Single sweep.

[20] Alder, Baldinger, Huber, and Metzger, Helv. Phys. Acta **19**, 207 (1946); **20**, 73 (1947).
[21] H. Saltzmann and C. G. Montgomery, Rev. Sci. Instr. **21**, 548 (1950).
[22] C. Balakrishman and J. D. Craggs, Proc. Phys. Soc. London **A63**, 358 (1950).
[23] L. B. Loeb, following paper [Phys. Rev. **97**, 275 (1954)].

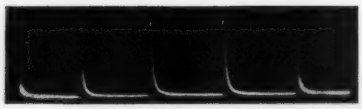


FIG. 10 Streamers in oxygen, 300 mm Hg, 3050 volts, 1000 μsec full scale, $RC=3$ μsec. Single sweep—gain 1/200 that on Fig. 9.

FIG. 11. Streamers in oxygen, 300 mm Hg, 3050 volts, 1 μsec full scale, $RC=0.003$ μsec. Many sweeps superposed.

statistical effect account for the fluctuations. Because of their limited spread, the fluctuations are smaller in amplitude than they were in air. These pulses were observed, at the appropriate V, at all pressures in pure oxygen, with smaller ones present even without the α-particle triggering. They represent an important phenomenon setting a lower limit to velocity of pulse propagation, as shown by Loeb.[23]

Figure 10 illustrates another type of pulse observed at still higher applied voltages. These were present at pressures of 200 mm and above, and required no α triggering for their initiation. They differed markedly from the burst pulses in several respects: Their amplitude was much higher (gain in Fig. 10 is less by a factor of 200 than in Fig. 9). They were spaced at uniform intervals of time and were approximately constant in amplitude at a given V and p. Their duration, as can be seen in Fig. 11 with a faster sweep speed, was a great deal shorter (10^{-7} sec as opposed to 10^{-4}). Instantaneous currents in these pulses were as high as 10^{-4} amp, and it is estimated (from the pulse area of Fig. 11) that they contain approximately 10^8 electrons.

These properties are characteristic of streamers, which like the burst pulses were originally observed in point-to-plane geometry[17,18] and were suspected to have been observed by Miller and Loeb.[3] These oscillograms and data for the first time consequently indicate the existence of streamers in Geiger counter geometry. The basic processes in streamer formation have been described by Loeb.[24] The discharge spreads in a series of steps very similar to those involved in the spread along the wire in air, except that for streamers the propagation is radially outward toward the cathode instead of along the axis. The outward direction for the propagation is favored by extremely localized photoionization, i.e., when the steps become very short.

It is interesting to note on comparison of Figs. 8 and 10 that streamer spacing is closely related to ion transit time. At all pressures where they were observed, streamers were separated by intervals equal to roughly ⅔ of the duration of the single ion pulse appearing at lower voltages. Ion transit times are shorter in the streamer region than for ion drift pulses, both because the applied voltage is higher and because of much denser space charge increasing cathode fields. It appears, therefore, that the ions in one streamer which are created closest to the wire must travel essentially all the way to the cathode before the field recovers sufficiently for a second streamer to begin. It is in fact possible that the arrival at the cathode of positive ions created in one streamer actually results in the liberation of secondary electrons which serve to trigger the next streamer as observed by A. F. Kip in point-to-plane corona with a screening gauze over the cathode.

[24] L. B. Loeb, *Fundamental Processes of Electrical Discharge in Gases* (John Wiley and Sons, Inc., New York, 1939).

ACKNOWLEDGMENTS

The writer wishes to thank Professor L. B. Loeb, who proposed this investigation, including the use of the transverse pulse, and under whose guidance this work was carried out, for his encouragement and suggestions. Special thanks are also due to Professor J. H. Reynolds and Mr. S. Balestrini for the mass spectrographic analyses, and to Mr. M. Corbett for his expert assistance in glassblowing.

Corona Threshold and the Range of Velocities of Pulse Spread in Geiger Counters*

LEONARD B. LOEB
Department of Physics, University of California, Berkeley, California
(Received October 5, 1954)

By using the writer's expression for a self-sustaining burst pulse corona and Alder, Baldinger, Huber, and Metzger's expression for the propagation of a Geiger pulse along the anode wire in coaxial cylindrical counter geometry, the limits of the velocity of spread of counter pulses can clearly be delineated. The interpretations resulting clarify the significance of the previously assumed fictive average distance of spread. These relations are illustrated in terms of the observations of E. Huber on O_2-N_2 mixtures.

IN 1948, the writer[1] set the threshold for the corona discharge in Geiger-counter-type geometry as

$$0.5 f_1 f \exp(-1.86 r_{0g}\mu) \exp\left(\int_{r_{0g}+a}^{a} \alpha dr\right)=1. \quad (1)$$

Here the 0.5 is a geometrical factor for loss of photons; f is the number of photoionizing photons of absorbtion coefficient μ created per avalanche electron; r_{0g} is the length of the zone for effective ionization, i.e., to start a full avalanche of $\exp(\int_{r_{0g}+a}^{a}\alpha dr)$ electrons at the anode wire of radius a; f_1 is the chance that an absorbed photon ionizes. The numerical factor 1.86 is one obtained by averaging the distance at which a photon must be created to yield an avalanche of radial length r_{0g}. In this relation nothing is said about propagation of the ionizing sequence down the wire with a velocity v. The condition cited marks the threshold of a self-sustaining photoionizing discharge which *may* spread down the wire yielding a Geiger counter pulse, or, *may* continue locally, whether it spreads or not, until space charge accumulation extinguishes the discharge.

In 1947, Alder, Baldinger, Huber, and Metzger[2] derived a relation for the spread of a Geiger-counter pulse with velocity v down the wire and set this as the threshold for the Geiger-counter action, since the pulse to be counted must come from the whole wire. It has been well known that the velocity of spread of Geiger-counter pulses is generally confined within narrow limits and is determined by composition of the gas, applied potential, and geometrical counter characteristics. The limitations of this range of velocities has not been clearly delineated. Very recently, Huber,[3] in the writer's laboratory, measured the velocity of spread of pulses in mixtures at constant potential under conditions which were simpler than the usual fast-counter mixtures. She noted, as well known, that the velocity decreased with increase in content of photoionizable gas at constant potential; in this case, the gas being O_2. It was also observed that at the upper limit of 40 percent O_2 there was no more α-particle triggering of a pulse spreading down the wire. There was, however, a definite discharge threshold producing an intermittent series of smaller discharges that did not spread down the wire. At low O_2 content the velocity of spread was high but higher potentials were needed to give pulses, and at some maximum velocity the Geiger-counter action gave place to a Townsend discharge at lower potential, there being no pulses.

The limitations on velocity can be clearly delineated if the relation of Alder, Baldinger, Huber, and Metzger, expressed in terms more correct as regards amplification factor in the Townsend avalanche, is compared with the writer's threshold relation. The modified relation reads

$$0.5 f_1 f_0 \frac{p_a}{p} \exp\left(\int_{r_{0g}+a}^{a} \alpha dr\left(\left[\exp\left(-\mu_0\theta v\frac{p_1}{p}\right)+\mu_0\theta v\frac{p_1}{p}\,\mathrm{Ei}\left(-\mu_0\theta v\frac{p_1}{p}\right)\right]\right)\right)=1. \quad (2)$$

Here all terms are the same as the writer's except that p_a is the partial pressure of the gas creating the ionizing photon, p_1 is the partial pressure of the gas being photoionized; p the total pressure, Ei stands for the exponential integral and $\theta v=x_0$ measures a critical distance along the anode wire needed to propagate a pulse with velocity v. Here θ is the longest time needed for the photon emission from the beginning of the avalanche. The partial pressure ratios should have been included in the writer's threshold relation for corona but were not included as they were implicitly contained in the definitions of f and μ which are; $f=f_0p_a/p$ and $\mu=\mu_0p_1/p$. The presence of the second term in the bracket comes from the averaging process for photon creation with x component in excess of x_ℓ along the wire. It fulfills in a sense the same function as the author's factor 1.86 but applies to the propagation concept. In general, it is not an essential term.

Thus the threshold for a self-sustaining corona and that for propagation of a counter pulse with velocity v down the wire differ in principle only in that one has a length $x_0=\theta v$ for propagation in the absorption term while the other uses in fact the critical ionizing avalanche distance $1.86r_{0g}$.

* The considerations follow studies carried out under grants from the U. S. Office of Naval Research.
[1] L. B. Loeb, Phys. Rev. **73**, 798 (1948).
[2] Alder, Baldinger, Huber, and Metzger, Helv. Phys. Acta **20**, 73 (1947).
[3] E. Huber, preceding paper [Phys. Rev. **97**, 267 (1954)].

Now the time θ is composed of the time for the creation of the avalanche, the time for photon travel, and the time of radiation. If $v(r)$ is the electron drift velocity as a function of distance from the anode, θ cannot be less than

$$\tau=\int_{r_{0g}+a}^{a}[1/v(r)]dr; \tag{3}$$

and in fact, if the active photon emission has states of average lifetime τ_f, then the θ defined by Alder, Baldinger, Huber, and Metzger becomes[3]

$$\theta=\int_{r_{0g}+a}^{a}\frac{dr}{v(r)}+\tau_f,$$

since the time of photon transit across 10^{-2} cm with the velocity of light is negligible. Usually the first term of θ is the important one though this need not always be true.

The relation $x_0/\theta=v$ has significance only in defining a critical distance needed for propagation of the pulse along the wire with a velocity v. Neither x_0 nor v is fixed to any basic physical process by the theory involved. In practice, v can be observed and θ is a time definitely fixed by physical processes. From θ and v, then, a distance x_0 can be calculated. However, as derived, x_0 and v may take any appropriate values as long as they maintain a ratio which usually depends on the field geometry, the applied potential, the variation of the Townsend coefficient α with X/p defining r_{0g} and the drift velocity of electrons.

There are, however, limitations imposed on the velocity of propagation v, both upper and lower, which can be discussed quite separately from the involved relations in Eq. (2). The upper limit to v is set by the writer's relation (1). This relation defines a minimum average distance r_{0g} of avalanche travel to the anode at the corona threshold V_g. This distance is determined by the field geometry, X as a function of r, the variation of the Townsend coefficient α/p with X/p, and thus with r, the partial pressures p_a, p_1, the total pressure p, and the probabilities f_0 and f_1. Below the average value of r_{0g}, avalanches cannot start a self-sustaining corona. At the threshold with r_{0g}, the corona may or may not propagate as a pulse down the wire. Since the time of avalanche creation and photon radiation is θ, and until these acts are carried out in a time θ, the avalanche effectiveness has not been achieved so that there is a maximum velocity of spread v_m defined by $\theta=r_{0g}/v_m$ which cannot be exceeded. In consequence, in general,

$$v\theta=x_0\leq r_{0g} \quad \text{and} \quad r_{0g}/\theta\geq x_0/\theta=v. \tag{4}$$

At v_m, $x_{0m}=r_{0g}$. At all other times it is less. Thus the velocity of propagation along the wire can never exceed the velocity of avalanche creation and completion of photoionization, a self-evident fact but not previously indicated.

There is also a lower limit, v_l to v and thus to x_0. If x_0 becomes very small the travel of the pulse will be so slow that a succession of avalanches can built up the local space charge so that the pulse is quenched before it has progressed very far. Especially near and just above threshold V_g, the succession of photoionizing processes will quickly build up a space charge capable of quenching the pulse. Assume that it takes only a succession of n avalanches to quench a corona by space charge. This number n may not be very great. A succession of some 10 photon-produced avalanches caused a marked diminution of Lauer's photon, γ_p pulses in H_2 though it did not quench them. At a few volts lower they would have done so. The time taken by n avalanches in succession is $n\theta$. Thus if

$$r_{0g}/v_l=n\theta=r_{0g}\theta/x_{0l}, \quad v_l=r_{0g}/n\theta, \tag{5}$$

there will at V_g be a threshold of corona discharge which will propagate a short way down the wire but will be quenched by space charge before it covers the whole anode. There will be, by longer-range photon absorption, new pulses started at different points on the wire, but they will extinguish locally. This will give an intermittent corona of many small localized pulses seen to differ on the oscilloscope screen from the larger sequence of heavy pulses where propagation down the wire occurs as noted by Huber.[3] Such phenomena were not noted by counter workers as they usually operated well above V_g. Thus again $v\geq r_{0g}/n\theta$ sets a limitation on the velocity and on the minimum value of

$$x_{0l}=v_l\theta=r_{0g}/n.$$

It may be of interest briefly to regard the conditions leading to upper and lower limits to v and x. Obviously these limits are governed by the very complex interrelations indicated by the threshold equations (1) and (2). They involve primarily the variation of θ, the applied V including α and X/p which depends on geometry, the nature of the gas, etc., including, of course, f_1, f, and μ, and the rate of variation of θ with V relative to r_{0g}. If, however, a gaseous mixture can be used in which V_g remains sensibly constant, or in which V can be maintained constant and with constant pressure and geometry r_{0g} remains constant, with θ relatively fixed, the situation simplifies. It happens that N_2-O_2 mixtures have relatively so little variation in this respect that it is possible to consider the only important factor that does vary the absorption coefficient $\mu=\mu_0 p_1/p$ which is entirely caused by the O_2, for which the partial pressure p_1 can be varied. With 1 percent O_2, $\mu\sim 0.5$ cm^{-1}; and at 100 percent O_2, $\mu\sim 50$ cm^{-1}. It is seen that if other things are to remain constant, the quantity $-\mu x_0$ in the exponential term must be maintained relatively constant for sensibly the same threshold. This means that at 1 percent O_2, x_0 must be large and at 100 percent O_2, x_0 must be

quite small. Huber observed that pulse spread velocity with 5 percent O_2 was $\sim 9\times 10^7$ cm/sec and at 40 percent O_2 pulses no longer spread, the velocity in air at 20 percent O_2 being 4×10^7 cm/sec. Thus at 5 percent O_2 with $\mu\sim 2.5$ cm^{-1} the velocity of spread was high, while in air at $\mu\sim 10$ cm^{-1} the velocity was lowered to about half. Thus a change in μ by a factor of 4 reduced v by a factor of 2, and at constant θ this means a factor of 2 change in x_0. When μ increased by a factor of 8 and in theory x_0 was reduced by the same amount, pulses ceased to propagate. Since, however, even with V constant, as V_θ varies with p_1 relative to V the constancy of the other variables was not sufficient to warrant accurate quantitative comparison. But it is seen that, as observed for counters, decrease in μ increases v.

Actually it is probable that even at 1 percent O_2 there will be pulses which spread down the wire, but these may require higher potentials than at V_θ since $r_{0\theta}/\theta$ at threshold may not be adequate or give the v required. When Huber changed the pressure of the active gas by changing the total pressure of air, the *increase* in μ at higher pressures appeared to *increase* v. However, there were accompanying changes in p, changes in V_θ, in $r_{0\theta}$, and in drift velocities of electrons, so that the reversal of variation with μ is not significant.

In principle, however, the preceding analysis has given an indication, separate from the very intricate threshold relations, as to physical limitations to be expected on the values of v and x_0, making these quantities more physically significant than the ratio of a fictive distance x_0 and a velocity v.

PHYSICAL REVIEW VOLUME 97, NUMBER 2 JANUARY 15, 1955

F Centers in Pure and Hydride-Containing Alkali Halide Crystals

R. S. Alger and R. D. Jordan
United States Naval Radiological Defense Laboratory, San Francisco, California
(Received July 7, 1954)

Additional evidence regarding the role of imperfections in the gamma-ray coloring and optical bleaching of the alkali halides has been obtained by comparing crystals with and without *U* centers. Parameters varied were gamma-ray exposure rate, crystal temperature, hydride concentration, *F*-band illumination intensity, and thermal history of the crystal. The coloring efficiency increased with temperature, trap depth and vacancy or *U*-center concentration. Under *F*-band illumination, both pure and *U*-centered crystals exhibited an initial period of rapid bleaching, followed by a long slow-bleaching period. Interrupting the bleaching in *U*-centered crystals by storage in the dark restored the higher bleaching rate. The slow bleaching was temperature-dependent, increased slightly with *F*-band illumination intensity, and was essentially independent of the *U*-center concentration. The results support the hypothesis that coloring and bleaching rates are primarily controlled by the availability of suitable holes and vacancies. Crystals containing the hydride color more efficiently than pure crystals because all constituents of the *F* center are present in the parent *U* center; whereas pure crystals bleach faster because of the accessibility of the holes.

I. INTRODUCTION

ALTHOUGH the coloring of alkali halide crystals by ionizing radiations is an old field of study, improvements in experimental technique have continued to yield valuable information concerning the photochemical reactions involved in the coloring process. Interest in the relations between the various crystal imperfections has been stimulated by recent efforts to combine the extensive knowledge of photochemical, photoelectric, and luminescent phenomena into a unified system of defects and reactions.[1]

The object of the current study is to compare the rate, yield and energy requirements for some of the photochemical reactions in pure KBr and in mixed crystals of KBr plus KH in order to gain further insight regarding the effect of the type and concentration of crystal defect on interactions with transient imperfections such as photons. One advantage in the study of *F* centers produced in mixed crystals rests in the fact that both the *F* centers produced and the number of *U* centers destroyed by the transient defects can be measured, thereby providing a check on electrons lost to undetected defects. Although mixed crystals have been used for most of the observations, a sufficient quantity of pure KBr has been included to permit comparison to the extensive data on pure crystals in the literature.

II. EXPERIMENTAL PROCEDURE

Because of the variety of measurements considered, specific experimental details are given in the appropriate section. Crystals grown by the Harshaw Chemical Company were filled with *U* centers by heating the crystals in a mixed atmosphere of potassium vapor and hydrogen. Samples to be compared were split from adjacent positions in the original crystal and given uniform

[1] F. Seitz, *Imperfections in Nearly Perfect Crystals* (John Wiley and Sons, Inc., New York, 1950), p. 13; Revs. Modern Phys. 26, 7 (1954).

thermal histories by simultaneous heat treatment. Uniform temperatures were obtained by heating the crystals in stainless steel chambers either embedded in a large copper block or submerged in a bath of molten aluminum. Radiation sources included x-rays from a 50-kv tungsten target (beryllium-window) tube and a 250-kv tungsten target tube and gamma rays from 0.5- and 10-curie Co^{60} sources. X-ray exposures were measured with air ionization, thimble-type chambers. The 0.5-curie Co source was calibrated by the National Bureau of Standards. All *F*- and *U*-center concentrations were determined from absorption measurements made with a Cary Model 11 recording spectrophotometer.

III. PHOTOCHEMICAL COLORING

A. Pure KBr

In discussing the factors controlling photosensitivity, it is advantageous to begin with crystals before the *U* centers have been added. All crystals contain imperfections in the form of vacancies (lattice positions where ions are missing, called Schottky defects) and dislocations. Equal numbers of positive and negative-ion vacancies are required to maintain electrical neutrality, and vacancies of opposite sign become attached to form pairs. The measurements of Delbecq *et al.*[2] indicate that essentially all Schottky defects are paired at room temperature in crystals of KBr and KI leaving few single vacancies available for *F*-center formation.

When high-energy photons undergo photoelectric, Compton or pair-production interactions in the crystal, the energetic electrons released produce many secondary ionizations and excitations. Some electrons form *F* centers by becoming trapped in vacancies. Others return to neutral bromine atoms (holes left behind in the ionizing process) and are removed from the picture as they reform bromine ions. Part of the absorbed energy presumably goes into the formation of single vacancies either through disruption of Schottky pairs

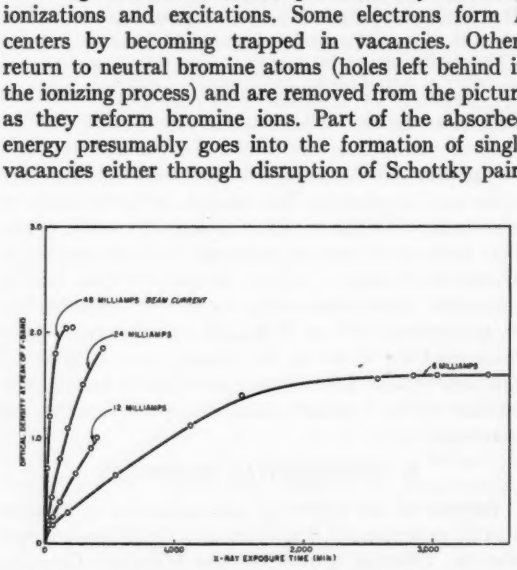


FIG. 1. *F* centers produced in KBr as a function of x-ray energy absorbed for four rates of irradiation.

[2] Delbecq, Pringsheim, and Yuster, J. Chem. Phys. 19, 577 (1951).

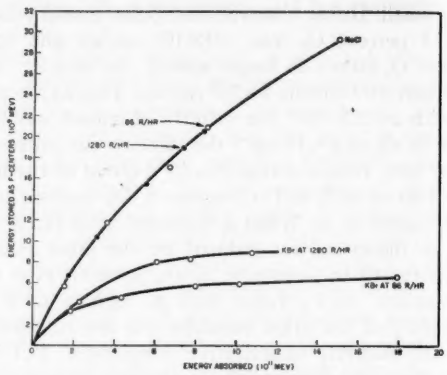


FIG. 2. Influence of energy absorption rate on *F*-center production rate and equilibrium concentration.

or by interactions between excitons and lattice dislocations. If the *F*-center concentration is plotted as a function of energy absorbed by the crystal, the coloring rate gradually decreases with the total energy until an equilibrium value is reached (Figs. 1 and 2). Presumably the electrons in the conduction band are captured initially by existing or easily formed single vacancies; however, these traps become depleted as the irradiation progresses, and the rate of vacancy formation becomes a governing factor in color center production. The coloring rate is further reduced by thermal bleaching and other forms of *F*-center excitation until at equilibrium, the rate of *F*-center production equals the rate of destruction. Apparently, both the coloring rate and the equilibrium concentration depend on the relative availability of single vacancies and competing holes, the stability of the *F* centers formed, and the rate of energy absorption (see Figs. 1, 2 and 5). Increasing the vacancy concentration and the probability of vacancy creation at dislocations by thermal and mechanical deformation of the lattice has long been known to increase the coloring ability. The most perfect natural alkali halide crystals which have had years to reach thermal equilibrium are difficult to color with x-rays or gamma rays; however, these crystals color easily after they have been heated near to the melting point and quenched.[3] At low coloring rates, it is possible for an *F*-center equilibrium concentration to exist before the scarcity of vacancies dictates a saturated condition. As the ionization rate increases, the loss of *F* centers through thermal bleaching becomes less important, and coloring is limited by the lack of vacancies. In KBr, thermal bleaching at room temperature is appreciable in exposures which last for several days; consequently, the equilibrium *F*-center concentration varies with the rate of energy absorption. In Fig. 1 an eight-fold increase in the x-ray beam current (50-kv) raised the equilibrium density by a factor of about 1.3 and similar

[3] Delbecq, Pringsheim, Vareck, and Yuster, U. S. Atomic Energy Commission Report AECU-1533, 1950 (unpublished).

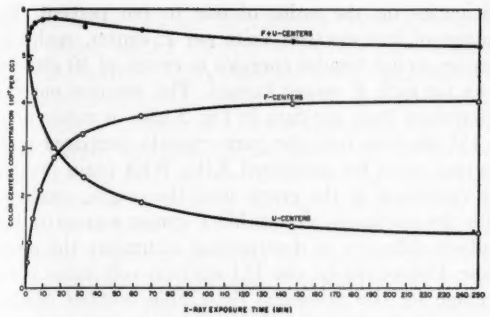


FIG. 3. X-ray-induced U- to F-center reaction in KBr.

results were obtained with Co^{60} for a fifteen-fold increase in energy-absorption rate (see Fig. 2). The initial coloring rate was about the same for all exposure rates used in Figs. 1 and 2; but, during the fairly linear portions of the curves, the lower exposure rates are less efficient for coloring; however, vacancy formation as well as thermal bleaching were probably contributing here. Since thermal bleaching is less severe in more tightly bound F-center electrons, the exposure rate has a smaller effect on NaCl than on KBr (see Fig. 2).

B. $U \rightarrow F$ Reaction

Processes occurring in crystals containing U centers are analogous to those just described; however, the concentration and characteristics of the vacancies and holes formed produce an appreciable enhancement in photosensitivity. If both the U and F bands are measured as a function of x-ray energy absorbed, it becomes apparent that most of the F centers are produced at the expense of U centers.[4] Figure 3 shows the F- and U-center concentrations, as well as their sum, plotted as a function of the x-ray energy absorbed. Initially, the total color-center concentration increases gradually with exposure, presumably because some F centers are created directly from the lattice; however, this contribution is less than ten percent. As the exposure continues, the total concentration decreases as electrons are lost to other centers not measured in the F and U bands. Good evidence exists for the two-step nature of the resonantly excited $U \rightarrow F$-center reaction in which first the electron is excited and second the neutral hydrogen migrates to an interstitial lattice position.[5] With ionizing radiations the hydrogen may lose an electron either through direct excitation or by hole conduction to an adjacent neutral bromine, thereby freeing the neutral hydrogen. Each U center represents a prospective F center, and a large U-center concentration has the same effect on photosensitivity as a large single-vacancy population. Figure 4 shows curves for crystals having varying U-center concentrations, but identical thermal histories produced by heating samples from a single original crystal through an identical thermal cycle. The variation in U-center concentration was obtained by limiting the amount of potassium metal added to the chamber in the coloring process. The reverse reaction of F center$\rightarrow U$ center proceeds slowly; consequently, a further gain in sensitivity results from reduced thermal bleaching, and the equilibrium color-center concentration for crystals containing U centers is much higher than for plain crystals as shown in Figs. 4 and 5.

C. Energy Storage Efficiency

F-center formation can be considered as an energy storage process in which about five electron volts are stored per color center in KCl and KBr. The storage efficiency serves as a figure of merit for photosensitivity. Quantitative energy measurements also can be used to determine the average energy required to convert a U center to an F center. Energy measurements were made on both pure and sensitized crystals, using Co^{60} gamma rays and an exposure rate of about 6.5 r per hour. The energy absorbed in one cm^3 of crystal was calculated on the basis of the following assumptions:

1. At the average gamma-ray energy of 1.27 Mev, the photoelectric absorption contributes only about one percent to the energy absorbed, and pair production is nil; consequently, only Compton interactions are considered.

2. All of the energy transferred to secondary electrons is absorbed in a crystal except for the electrons originating near the crystal surface. A correction factor of 0.87 was included to account for the calculated loss of 13 percent of the energy carried through the crystal surfaces by escaping electrons.

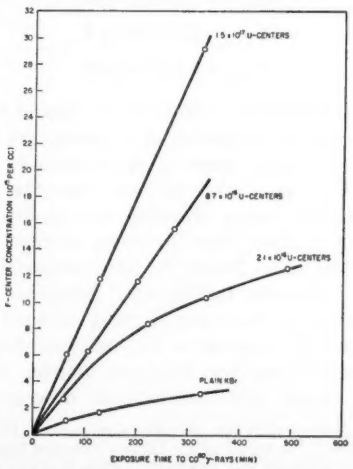

1.5 × 10^17 U-CENTERS
8.7 × 10^16 U-CENTERS
2.1 × 10^16 U-CENTERS
PLAIN KBr
F-CENTER CONCENTRATION (10^17 PER CC)
EXPOSURE TIME TO Co^60 γ-RAYS (MIN)

FIG. 4. Effect of U-center concentration on F-center production rate in KBr under gamma-ray bombardment.

[4] H. Pick, *Photographic Sensitivity* (Butterworths Scientific Publications, London, 1951), p. 9.

[5] N. F. Mott and R. W. Gurney, *Electronic Processes in Ionic Crystals* (Oxford University Press, London, 1948), second edition; Hans Tomine, Ann. Physik 5, 601 (1949).

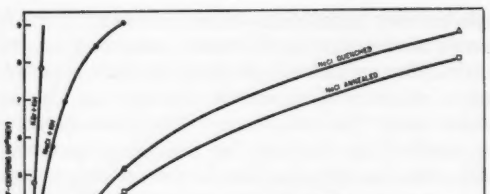


FIG. 5. Effect of crystal material and past thermal history on energy storage efficiency.

3. This correction factor was calculated assuming that all secondary electrons have the average energy of Compton-scattered electrons and the direction of travel corresponds to the average energy. If the correction is made for this surface loss neglecting the small number of electrons originating near the edges, the energy absorbed per gamma ray is 0.075, 0.056, and 0.063 Mev for one cm³ of KBr, KCl, and NaCl, respectively.

The energy absorbed per unit volume of crystal is

$$E/V = N\sigma\epsilon/V,$$

where ϵ = the average energy of a Compton electron produced by Co^{60} gamma rays = 0.613 Mev, σ = the total Compton cross section for Co^{60} gamma rays, N = number of incident photons, and V = crystal volume = 1 cm³. Since the Co^{60} source was suspended in mid-air, a reasonably pure spectrum was obtained. Optical measurements were made on the Cary spectrophotometer, and the F-center concentrations calculated using the following equations[6]:

$$N_{NaCl} = 4.8\times10^{15}\times K,$$
$$N_{KCl} = 3.7\times10^{15}\times K,$$
$$N_{KBr} = 3.6\times10^{15}\times K,$$

where K = absorption coefficient at F-band maximum in cm^{-1}, and N_{salt} = number of F centers per cc. From the plot of energy stored as a function of absorbed energy (Fig. 5), it can be seen that the efficiency (energy stored/energy absorbed) at the beginning of irradiation is about four to seven percent for plain crystals, and about 20 percent for sensitized KBr. This efficiency decreases rapidly with increasing F-center concentration in the pure crystals as equilibrium is approached. In a survey of the literature dealing with photochemical reactions involving ionizing radiations, Frerichs[7] found that the average process leads to efficiencies on the order of one to ten percent. For storage of five electron volts per F center, such efficiencies would require energies in excess of 50 electron volts for each F center formed. The reaction energies determined from the data in Fig. 5 lead to values of 92 to 154 electron volts for pure crystals compared to 26 electron volts for sensitized KBr. With plain crystals the curvature in the graph near the origin, combined with the minimum detectable F-center concentration, leads to difficulty in determining accurately the initial slope. Consequently, the 154 electron-volt value is the average for the minimum detectable number of electrons, and the 92 electron volts is for an estimated slope at the origin. These results are somewhat lower than those of Pick who obtained 35 electron volts for sensitized KBr using 50- to 60-kv x-rays. The Co^{60} determination of the reaction energy is probably low due to the assumption made in calculating the energy absorbed. The calculation assumed that six-tenths of the gamma-ray energy absorbed in the Compton process was taken by the Compton electron and the remaining energy was carried out of the crystal by the degraded photons. Many of these degraded photons will undergo additional photoelectric or Compton interactions; thus, the calculated energy will be too low by the amount of energy absorbed in the second collisions. The difficulty in correcting for secondary collisions makes a completely experimental determination of reaction energy more desirable, and such an experiment is in progress utilizing 2-Mev cathode rays from a Van de Graaff generator.

The choice of crystal material offers one method of influencing the energy storage efficiency. When the coloring rates in cubes of NaCl, KCl, and KBr are compared as in Fig. 5, the initial slopes are observed to be about the same for all three pure materials; however, as equilibrium is approached, the curves diverge and the materials having more stable F centers reach a higher equilibrium F-center concentration. The F centers of NaCl, KCl, and KBr require 2.63, 2.20, and 1.96 electron volts, respectively, for excitation at room temperature; consequently, the increased stability of NaCl F centers would contribute to the higher equilibrium value, but it must be remembered that the rate at which vacancies become available, also strongly influences this portion of the curve. By varying the two factors of stability and vacancy availability through the introduction of U centers, it is possible to override the influences of crystal material. Figure 5 also shows the coloring for the KBr and NaCl containing U centers. It will be seen that sensitized KBr, in addition to averaging about ten times the sensitivity of the plain material, rises to a higher equilibrium value than the NaCl. At present the highest U-center concentrations have been obtained in KBr.

D. Temperature Dependence

The thermal effects are negligible for the primary interactions because of the high energies involved; how-

[6] R. W. Pohl, Proc. Phys. Soc. 49, extra part, 3 (1937).

[7] R. Frerichs, "A Study of the Older Literature Related to the Coloring of Crystals by Irradiation with X-Rays or Nuclear Radiation," Northwestern University report, September, 1950 (unpublished).

ever, the mobilities and diffusions of vacancies and the atoms which control coloring rates vary exponentially with temperature. Measurements at low temperatures are complicated by the appearance of new traps which are thermally unstable at room temperature, e.g., the F' and V bands; however, certain trends can be detected in the existing data. Harten[8] measured the temperature dependence of the coloring rate in plain KCl over the range from −180°C to 60°C. This measurement showed the following general features: (a) The general coloring ability reached a maximum between −20°C and −75°C; and (b) The initial slope of the curve increased with temperature from −180°C to −20°C and remained constant from that point on to 60°C. The equilibrium F-center concentration decreased for all temperatures above 75°C, and since the curves at −110°C and −180°C did not reach equilibrium, the lower temperature limit for this trend cannot be fixed from this data. Duerig[9] measured coloring rates at 5° and 78° Kelvin, and found a higher rate at 5°, which suggests that the higher mobility at 78° Kelvin may cause more F centers to be bleached, thus reducing the coloring rate. If the initial slope depends on vacancies that either exist or are readily available, increasing the thermal agitation would assist in production of sites for F-center formation; therefore, the initial slope of the coloring curves should increase with temperature until thermal bleaching effects begin to dominate. The equilibrium value on the other hand would tend to increase as the temperature is lowered until the low temperature traps begin to dominate.

Tomine[5] studied the U-band absorption of KBr+KH throughout the temperature range from −253°C to 500°C and found that the quantum yield, i.e. (F centers produced)/(uv light quantum absorbed), had a value of unity at 500° and decreased to 20 percent at 20°, becoming zero at −120°C. Since U-band photons are resonantly absorbed by U centers, the thermal behavior presumably results from changes in the hydrogen atom mobility; therefore, Tomine's results substantiate the previously discussed two-step nature of the U-to-F conversion process as suggested by Mott and Gurney.[5] The possibility of F-center formation then becomes a temperature-dependent reaction which is quite sensitive to changes in the hydrogen mobility. At low temperatures, i.e., below −120°C, the details of the coloring picture are complicated by the formation of U' and U'' centers. In sensitized crystals colored with x-rays, the thermal effects have a possibility of modifying several stages in the process. The probabilities for Compton and photoelectric interaction processes are proportional to the first and fifth powers of the atomic number, respectively; consequently, Br^- with 36 electrons, compared to 2 for H^-, will monopolize the initial interactions. If direct ionizations or excitations by photoelectrons involve primarily the valence electrons, the fraction of the energy absorbed by the U centers would not account for the number destroyed. Crystals containing 10^{18} U centers per cm^3 have 10^4 Br ions present for each H^-; consequently, less than 0.01 percent of the absorbed energy is available for the $U \rightarrow F$ reaction instead of the 15 to 20 percent observed stored as F centers. Some mechanism such as migrating holes or excitons is required for transferring absorbed energy from Br^- to the U centers. This additional step offers another opportunity for the temperature dependence although the reduced hydrogen mobility is still probably the predominant thermal effect. Once the U center has received sufficient energy to excite the electron or the electron has been captured by a near-by hole, it is necessary for the hydrogen to depart as in the case of resonance absorption before a vacancy is available for F-center formation.

Figure 6 shows the result of measurements made on both plain and sensitized KBr samples exposed to x-rays over the temperature range from −180°C to 89°C. Sensitized crystals containing 10^{17} U centers per cm^3 exhibited the same temperature dependence as those containing more than 10^{18} U centers. The crystal temperature was maintained throughout the x-ray exposure by enclosing the samples in a copper chamber which in turn was immersed in the heating or cooling liquid. The filtration of the x-ray tube, plus that due to the crystal holder and Dewar flask, produced an effective x-ray energy of about 170 kv. The maximum color-center yield for plain crystals occurred between −60°C and 20°C, decreasing at both higher and lower temperatures in agreement with the work of Harten.[8] At elevated temperatures, the drop for plain crystals is due presumably to thermal bleaching; whereas the increased stability of sensitized crystals permits the yield to continue increasing throughout the temperature range covered. The reduction in the U band shown at

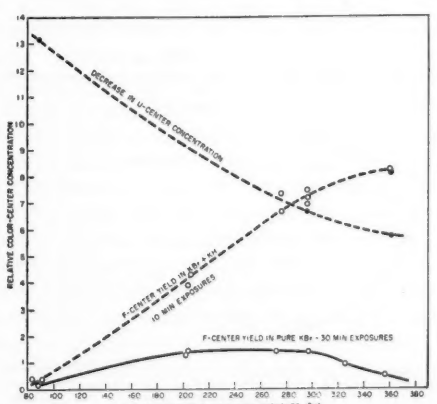

DECREASE IN U-CENTER CONCENTRATION
F-CENTER YIELD IN KBr + KH
10 MIN EXPOSURES
F-CENTER YIELD IN PURE KBr - 30 MIN EXPOSURES
RELATIVE COLOR-CENTER CONCENTRATION
CRYSTAL TEMPERATURE DURING X-RAY EXPOSURE (°K)

FIG. 6. Color-center yield in plain and sensitized KBr as a function of exposure temperature for x-ray exposures of 10 and 30 minutes, respectively.

[8] H. Harten, Z. Physik **126**, 619 (1949).
[9] W. H. Duerig, Ph.D. Thesis at University of Maryland, 1952 (unpublished).

the top of the graph agrees with the *F*-center production. All crystals were measured at room temperature after exposure to the x-rays.

E. Luminescence

When both pure and sensitized KBr, colored at liquid nitrogen temperature, are heated to room temperature, a small fraction of the stored energy is released as thermoluminescence; however, the magnitude and temperature location of the emission peaks do not coincide in the two materials. Table I lists the temperature location and relative intensity of the peaks between −196°C and 250°C observed at a constant heating rate of 21°C per minute. No emission between 2000 and 8000 A could be detected with 1P28 photomultiplier tubes for KBr+KH heated from 20°C to 350°C although complete bleaching occurred at the highest temperature. Since all of the crystals tabulated in Table I received the same x-ray exposure, the *F*-center concentration was an order of magnitude higher in the sensitized salt; therefore, radiationless transitions in KBr+KH account for a larger fraction of the thermal bleaching than in pure KBr. The extremely low quantum yield in pure KBr (approximately one photon per 100 *F* centers destroyed)[10] leads to the assumption that defects other than *F* centers and neutral bromine atoms are involved in the transition. The absence of thermoluminescence above room temperature and the low relative yield at low temperatures would suggest that KBr+KH either bleaches by a different process or cannot utilize the defects responsible for thermoluminescence in the pure crystals. The bleaching process will be discussed in subsequent paragraphs.

F. Effect of Past History

Previous mention has been made of the increase in sensitivity that can be obtained by mechanically or thermally increasing the vacancy concentration in plain crystals. Figure 5 shows coloring curves for KBr, KCl, and NaCl crystals both annealed and thermally shocked by quenching in air from 700°C to room temperature. The initial coloring rate is higher for all quenched crystals, and the knee marking the onset of equilibrium occurs later, indicating that a larger number of single or prospective vacancies are available in the heat-treated crystals. The effects of mechanical strain are usually apparent in crystals which have been split from larger chunks since the coloration is noticeably darker where some straining and cracking has accompanied the splitting. This effect of mechanical strain was not found in the sensitized materials, apparently due to the predominance of the *U*- to *F*-center transition which completely masked any additional sensitivity arising from mechanical factors.

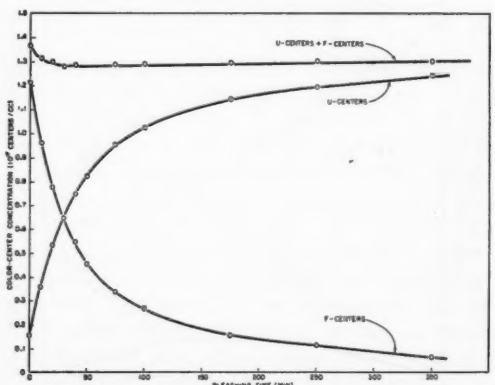


FIG. 7. Conversion of *F* to *U* centers in KBr by illumination in the *F* band.

TABLE I. Thermoluminescent peaks in pure and sensitized KBr.

Material	Temp. location of peak	Relative intensity
Pure KBr	−120°C	200
	− 80°C	50
	− 40°C	50
	0°C	50
	20°C	100
	60°C	50
Sensitized KBr	−140°C	1¼
	−130°C	1
	− 60°C	25

IV. PHOTOCHEMICAL BLEACHING

A. Pure KBr

Once again some insight into the mechanisms occurring in sensitized crystals can be obtained from the examination of data available for pure alkali halides. *F* centers can be destroyed and the crystal returned to its original colorless condition by optical, thermal or mechanical bleaching. During the bleaching process, some of the stored energy is liberated either as a direct consequence of the bleaching or as a byproduct in the form of luminescent radiation which has been termed photoluminescence, thermoluminescence, and triboluminescence, depending on the method of bleaching. The most outstanding feature of the bleaching process is the apparent difference in stability of the various color centers in a single crystal. Under constant illumination in the *F* band, the initial bleaching is rapid; whereas the final bleaching proceeds slowly and is sensitive to the origin, past history, and impurity condition of the crystal. Casler, Pringsheim, and Yuster[11] have observed that both the type of irradiation used and the coloring rate influence the *F*-center stability. For example, stability increased in going from beta-ray to soft x-ray to gamma-ray exposures, and decreased

[10] M. L. Katz, Izvest. Akad. Nauk (S.S.R.R.) **13**, 149 (1949).

[11] Casler, Pringsheim, and Yuster, J. Chem. Phys. **18**, 887 (1950).

with increasing coloring rate. Alexander and Schneider[12] observed an increase in F-center stability with time between coloring and bleaching. In the bleaching process, F centers escape from vacancies and return to holes, reforming the halogen ions; consequently, the V centers which result from the presence of neutral halogen atoms in the lattice are destroyed along with the F centers. The type of deficiency center, i.e., V_1, V_2, V_3, V_4, or alpha or beta, depends strongly on the irradiation temperature, and a corresponding temperature dependence of the F-center stability is observed.

Since the F bands do not change with the shifts in stability, the bleaching mechanisms proposed generally control the rate by availability of the holes.[11–13] Schneider[13] suggests that electrons and positive holes formed in the photochemical process remain in the same vicinity in the crystal, the electron as an F center and the hole either free or trapped as a V center. Rapid bleaching results from the recombination of electrons and holes that are close at hand; whereas the slow bleaching requires a disruption of more stable V centers or involves more remote holes and electrons. It will be observed that all of the bleaching-rate factors control in part the availability of receptors. For example, soft x-rays and beta rays produce very intense local ionization; consequently, the distance between F centers and holes will be less than in a crystal irradiated with higher energy gamma rays. Similarly, a storage period after irradiation or a low irradiation rate permits time for the more stable V-center complexes to form.

B. $F \rightarrow U$ Reaction

The distinguishing difference between the bleaching process in pure and sensitized crystals is the high stability of the F centers formed from U centers, which

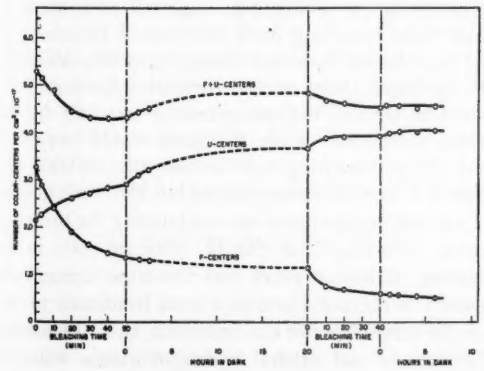


FIG. 8. Restoration of rapid bleaching rate in KBr+KH by periods of storage in the dark. The bleaching light intensity was approximately 5000 ft candles of white light. (Note the difference in time scales.)

[12] J. Alexander and E. Schneider, Nature 164, 653 (1949).
[13] E. Schneider, *Photographic Sensitivity* (Butterworths Scientific Publication, London, 1952), p. 21.

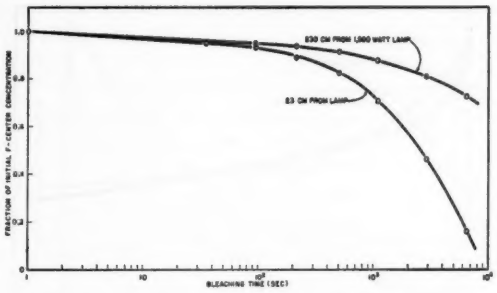


FIG. 9. Bleaching of F band in KBr+KH for a 100-fold variation in illumination intensity.

presumably is due to the difference in the behavior of the holes formed in the two processes. Figure 7 shows the typical behavior of the F and U bands under constant illumination from a tungsten filament lamp. The F band bleaches rapidly at first as in the case of pure crystals and the sum of $U+F$ centers decreases slightly, indicating that some electrons are returning to halogen neutrals. If bleaching is interrupted with periods of storage in the dark, the bleaching rate is observed to increase after each dark period as shown in Fig. 8. In some crystals the absorption increases on the long wavelength side of the F band, presumably due to F' or other centers which bleach out thermally during the dark storage period, giving rise to additional U centers. Although this increase in the bleaching rate is most pronounced during the initial rapid change in slope, the increase can be observed after each dark storage as long as F centers remain in the crystal.

C. Effect of Illumination Intensity

The availability of electrons for U-center formation depends on the electron concentration in the conduction band, which in turn is determined by the intensity of illumination in the F band. At sufficiently low light levels, the bleaching rate is controlled by the rate of electron release; however, at intensities corresponding to room light or brighter, swarms of electrons are available in the conduction band and other factors limit the bleaching rate. Figures 9 and 10 are typical examples of bleaching as a function of light intensity for continuous and chopped irradiation, respectively. In all cases, the samples compared were split from a single sensitized crystal and the illumination came from a 1500-watt tungsten lamp filtered to remove the infrared and supplying approximately 2000 ft candles at the 23-cm position. A one hundred-fold decrease in intensity for continuous illumination (Fig. 9) reduced the bleaching rate by a factor of 3 to 10. Similar curves for the average light transmitted by rotating sector disks show little change in slope after the first 100 minutes for a 20 to 1 variation in the number of holes in the chopper disk. The transition of electrons from one vacancy to another as a result of F-center destruction, plus re-

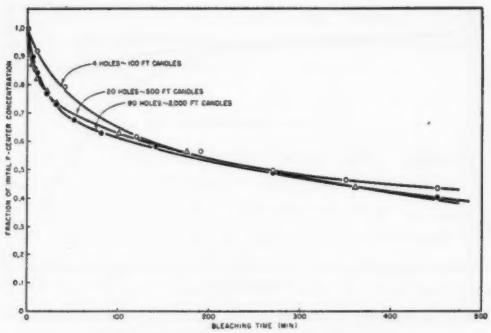

4 HOLES~100 FT CANDLES
20 HOLES~500 FT CANDLES
80 HOLES~2,000 FT CANDLES
FRACTION OF INITIAL F-CENTER CONCENTRATION
BLEACHING TIME (MIN)

FIG. 10. Bleaching of F band in KBr+KH for three different average light intensities as obtained with sector disks.

capture of an electron from the conduction band, can be looked upon as a form of temporary vacancy migration which could assist the bleaching process by bringing an unoccupied vacancy within reach of a hydrogen atom; consequently, the illumination effects are not solely tied to electronic transitions.

D. Thermal Effects

The F centers in sensitized crystals are quite stable up to temperatures of 100°C; however, the optical bleaching rate is strongly temperature-dependent even below this value. Figure 11 shows typical bleaching curves for three adjacent samples from a single sensitized crystal bleached at −196°C, 20°C, and 100°C. The −196°C and 20°C temperatures were maintained throughout the six hours of measurement while the 100°C crystal was cooled to room temperature for each F-center determination. At −196°C the F-center concentration dropped eight percent during the first 25 minutes and then remained constant. The crystal at 100°C was completely bleached in 50 minutes. Since the light intensity (approximately 2000 foot candles of white light) was the same as the maximum in Fig. 9, a large surplus of electrons was present in the conduction

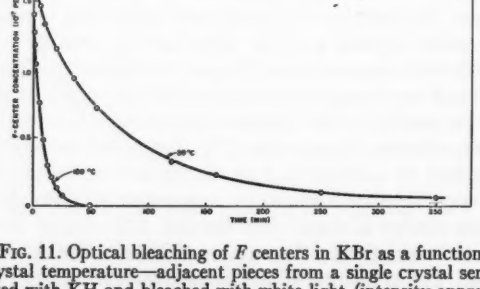


FIG. 11. Optical bleaching of F centers in KBr as a function of crystal temperature—adjacent pieces from a single crystal sensitized with KH and bleached with white light (intensity approximately 2000 ft candles).

band at all temperatures. Consequently, the thermal effect must involve other components of the U centers.

E. Effects of U-Center Concentration

Other factors which influence the bleaching process are U-center concentration, crystal thermal history, foreign ions, and crystal material. The study of U-center concentration effects is complicated by the difficulty in holding all other factors constant in preparing and bleaching the crystal. In order to assure a uniform past thermal history, all crystals compared here were heated simultaneously in the molten aluminum temperature bath using stainless steel chambers. Variations in U-center concentration were achieved by limiting the amount of alkali metal added to the containers. The pure crystals used in the comparison were put through the same heating and quenching cycles. Figure 12 shows the bleaching characteristics of crystals containing 0, 10^{16}, and 10^{18} U centers per cm^3 prior to the x-ray exposure. The abscissa has been related to the energy absorbed in bleaching the F band by graphically integrating the area under the time bleaching curve, allowing for the change in energy absorbed from the bleaching light as the absorption coefficient decreases with the reduction in F-center concentration. The pronounced difference in the bleaching rate between pure and mixed crystals is indicative of the relative ease with which F-center electrons can recombine with bromine and hydrogen atoms. In the pure crystal, bleaching occurs through capture by bromine atoms, while in crystals containing large U-center concentrations, essentially all of the F centers are formed through the disruption of U centers, and bleaching necessitates the recombination of the F-center electron with a hydrogen atom. Between these extremes are crystals containing small U-center concentrations in which case F centers will be formed from both H^- and Br^-; consequently, the bleaching curve would be expected to indicate an initial rapid bleaching until the neutral bromines are used up, followed by a slow bleaching as recombination with hydrogen atoms occurs. Increasing the number of U centers beyond the concentration required for 100 percent production of the F centers would have little effect on the bleaching rate, because two crystals with different U-center concentrations but the same number of F centers would appear identical during the bleaching process, assuming that the H^- play no part in the bleaching. If both crystals had the same thermal and impurity background and had been irradiated to produce the same F-center concentration, then the number of F centers and neutral hydrogen atoms would be identical and the recombination process should be identical. The experimental evidence lends qualitative support to this picture and the bleaching time is found to be of the same order of magnitude for all crystals starting with the same F-center concentration, but ranging in U-center concentrations from 10^{17} to 10^{18}

U centers per cm³. However, the difficulty in obtaining finer agreement would indicate that all of the other variables are not being held constant during the observation. Crystals containing 10^{16} or less *U* centers exhibit the initial rapid bleaching for a time comparable to that required to bleach the pure KBr followed by long, slow bleaching as shown in Fig. 12. Previous mention has been made of the fact that pure crystals that are hard to color are also difficult to bleach. Similarly in the mixed crystals, the more easily colored samples bleach faster.

F. Hydrogen Mobility Effects

If, as has been suggested, the bleaching rate in sensitized crystals is ultimately controlled by the behavior of the hydrogen, it might be possible to reduce the bleaching rate by forming *U* centers from deuterium. The quantum efficiencies for the photochemical decomposition of KH and KD molecules have been measured in KBr over the region from −100°C to 50°C. Throughout this region the quantum yield is about five percent lower for KD, suggesting that the deuterium has a lower diffusion coefficient.[14]

The crystals for this study were prepared by both the one- and two-step process. In the one-step process, samples split from adjacent portions of a single crystal were heated in identical chambers containing the same amount of potassium metal but under hydrogen and deuterium atmospheres. In the two-step process, the crystals were colored additively, then split and hydrogenated and deuteriumated in identical chambers. All heating was carried out in the molten aluminum bath.

Typical bleaching curves for crystals, sensitized with hydrogen and deuterium, are shown in Fig. 13. In all cases where comparable amounts of KH and KD were present and where similar *F*-center concentrations resulted from a given x-ray exposure, the over-all bleaching rate was higher in the crystals containing hydrogen. The initial rate in Fig. 13 was higher for KBr+KH; however, in other cases where both crystals had the same initial rate, the total bleaching time was longer in KBr+KD. The results lend qualitative support to the dependence of bleaching rate on hydrogen mobility, but improved crystal-preparation techniques are needed for a quantitative determination.

V. DISCUSSION AND CONCLUSION

In presenting the results on photochemical coloring, considerable attention has been given to the role played by holes and vacancies in the coloring process. The comparison between pure and sensitized KBr indicates two primary reasons for the enhanced photosensitivity in the mixed crystals; first, the *U* centers represent the storage of a large number of prospective vacancies which

[14] N. F. Mott and R. W. Gurney, *Electronic Processes in Ionic Crystals* (Oxford University Press, London, 1948), second edition, p. 149.

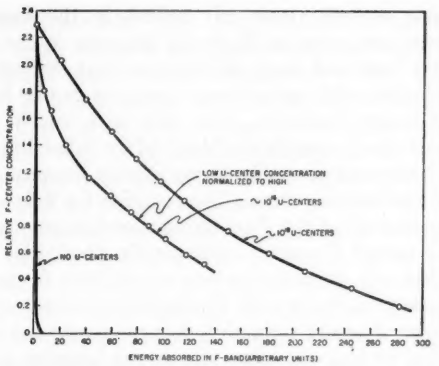


FIG. 12. Influence of *U*-center concentration on bleaching rate. Initial *F*-center concentrations in the ratio of 1 to 0.22 to 0.13 have been normalized. Crystals with low *U*-center concentrations received longer x-ray exposures.

can be utilized in *F*-center formation; secondly, the neutral atoms formed in the disruption of *U* centers offer little competition to the vacancies for the available electrons. Consequently, the color centers formed are more stable. If only the constituents involved in the coloring reaction are considered, the process in pure and sensitized KBr, respectively, can be symbolized as follows:

(a) *Pure* KBr:

$$Br^- + h\nu \rightarrow Br + e,$$

$$e + \text{vacancy} \rightarrow F \text{ center}.$$

(b) KBr+KH:

$$U \text{ center} + h\nu \rightarrow H + e + \text{vacancy},$$

$$e + \text{vacancy} \rightarrow F \text{ center},$$

where Br^- and Br represent the halogen ion and neutral atom, respectively; $h\nu$=photon energy; e=the electron; and, H=neutral hydrogen. In the pure crystal, the deionization of Br^- provides an electron plus a neutral atom. The vacancy required for *F*-center formation must either exist as a defect in the crystal or be created by another ionizing or exciting event.[15] The

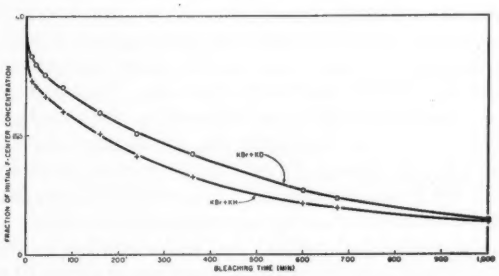


FIG. 13. Comparison of *F*-center bleaching in mixed crystals of KBr+KH and KBr+KD.

[15] F. Seitz, reference 1, p. 57.

bromine neutral atoms left behind in the ionizing process have a strong affinity for electrons in the conduction band and many electrons are lost through recombination with holes before a vacancy can be found for *F*-center formation. The ease with which pure crystals can be completely bleached by *F*-band illumination demonstrates this strong affinity between electrons and bromine atoms. The equation for KBr+KH shows that all of the *F*-center components are present in the parent *U* center; consequently, the initial high coloring rate continues as long as sufficient *U* centers are present in the crystal. The neutral hydrogen atoms left after *F*-center formation do not occupy the same position or play the same role as the bromine atoms produced in the pure crystals. The hydrogen occupies interstitial lattice positions and produces no potential disturbance that has been detected by the optical absorption bands; whereas the bromine atoms are confined to lattice sites and give rise to the various *V* bands. Initially, both pure and mixed crystals contain a concentration of vacancies which is determined by the past thermal history of the crystal. As coloring continues, these vacancies are used up by *F*-center production in pure crystals; but, when the *F* centers are produced from *U* centers, no change in vacancy concentration is required. The stability of *F* centers in the absence of *F*-band illumination gives some indication of the ease with which hydrogen and bromine atoms can recombine directly with *F* centers. Pure KBr crystals stored in complete darkness have been observed to lose one-third of the *F* centers because of thermal bleaching at room temperature over a period of a week's time; whereas mixed KBr+KH exhibited no detectable change in *F*-center concentration during the same time under identical storage conditions. A phosphorescence which can be observed visually in pure KBr continues for hours after extensive coloring by x-rays. KBr+KH shows no phosphorescence under identical conditions. Since the luminescence in the alkali halides accompanies the destruction of *F*-centers, this phosphorescent behavior of pure and mixed crystals is in accord with the observed stability of the *F* centers of the two materials in the absence of illumination. A neutral bromine atom located at a lattice site results in a hole in the filled band. This hole can migrate through the crystal by electronic transitions in which an electron jumping from a neighboring negative ion to the neutral atom appears as a hole migrating in the reverse direction. When through hole migration, a Br atom is formed in the vicinity of an *F* center, it is conceivable that the *F*-center electron could tunnel through the potential barrier to recombine with the bromine. Schneider[14] has concluded from data on the maximum *F*-center concentrations obtainable by the photochemical method that recombination by this tunneling process occurs for separations of about 10 to 30 lattice spacings. The high *F*-center concentrations and the stability of the centers formed in KBr+KH indicate a very low probability for the combination of *F*-center electrons and hydrogen atoms located at interstitial lattice positions.

For simplicity the coloring reactions were represented by direct interaction of photons with Br ions and *U* centers; however, most of the actual excitations occur by secondary processes, and a majority of the absorbed energy is dissipated in the crystal as heat, or lost through photoemission and is not stored in the form of color centers. A comparison of the energies required to produce the average *F* center in pure and sensitized crystals provides a more quantitative measurement of the efficiency of the reaction. Unfortunately, a wide variation in the conversion energy has been reported by the various investigators. Martienssen and Pick[16] found the initial rate to be constant in KBr+KH for all *U*-center concentrations above 10^{17} per cm^3, indicating that good energy transfer could occur up to distances of the order of 50 lattice spacings. Using 50- to 60-kv x-rays on crystals containing more than 10^{17} *U* centers per cm^3, Pick found that initially 35 electron volts were required per *F* center formed as compared with something in excess of 150 electron volts for pure crystals under similar conditions. Schneider[14] colored pure NaCl crystals with 40- to 50-kv filtered x-rays, obtaining an energy of less than 18 electron volts per *F* center formed, while Harten[8] found 120 electron volts were required per *F* center produced by x-rays in pure KCl. The initial coloring rates in pure KBr, KCl, and NaCl are quite similar as shown by Fig. 5. Consequently, the ionization energies found by Schneider and Harten should be comparable to the determination for pure KBr. The results discussed in Sec. III for coloring with Co^{60} gamma rays gave 26 electron volts per *F* center in sensitized KBr, compared to 92 to 154 electron volts in the pure crystal. Although the ionization energy determined with Co^{60} is probably on the low side, the initial ratio of about 4 between pure and sensitized crystals is valid and provides a comparison of reaction rates. Experiments now in progress, using 2-Mev monoenergetic cathode rays, should provide shortly a better figure for the reaction energy for irradiation in the Mev region.

The bleaching reactions in pure and mixed KBr crystals can be analyzed with reference to the following equations:

(a) *Pure* KBr:

$$F \text{ center} + h\nu \rightarrow e + \text{vacancy},$$
$$e + \text{hole} \rightarrow \text{Br}^{-};$$

(b) KBr+KH:

$$F \text{ center} + h\nu \rightarrow e + \text{vacancy},$$
$$e + \text{vacancy} + \text{H} \rightarrow U \text{ center}.$$

In the coloring process one hole was produced for each *F* center formed in the pure crystal while one neutral

[16] W. Martienssen and H. Pick, Z. Physik 135, 309 (1953).

hydrogen was formed for each F center in the sensitized KBr. Since the action of F-band illumination on an F center is identical in the two materials, any difference in the bleaching reaction must result from the processes represented by the second lines in the above equations. Schneider[14] has analyzed bleaching curves for pure salt in terms of exponentials and finds that a satisfactory fit can be obtained with two exponentials having decay constants of the order of 1.7×10^{-3} seconds^{-1} and 2.5×10^{-4} seconds^{-1}, respectively. He accounts for the two bleaching rates by the configuration of the hole involved in the reaction. In the fast decay, electrons released from F centers combine with nearby available single holes; whereas the slow decay is attributed to F centers which lose their electrons to bound holes such as V centers. About 50 to 70 percent of the centers were of the slow-bleaching variety. An analysis of bleaching data associated with Sec. IV shows qualitative agreement with Schneider's results although the experimental conditions were quite different, i.e., the F-center concentration was lower by a factor of two and the bleaching rate higher by an order of magnitude than reported by Schneider's experiment. The fast components decay at similar rates; whereas the various slopes obtained for the slow decays indicate more dependence on the past history of the crystal. The initial number of slow centers was never observed to exceed 40 percent and not all the bleaching curves for the pure crystals could be fitted by the above analysis, although there was no known significant difference in past history of the crystals.

The bleaching process in sensitized KBr involves the combination of three components to form the U center. There are four conceivable modes of the combination; first, a three-body collision; second, a direct recombination between the F center and a hydrogen atom; third, a hydrogen atom becomes trapped in a vacancy and electron capture follows; and fourth, an F-center electron attaches itself to a hydrogen atom producing an H^- which is captured by a vacancy. Mode 1 is statistically unlikely and Mode 2 has been ruled out by the stability of the color centers in the dark. The other modes depend on the rate at which the hydrogen ions and the vacancies come together. Qualitatively, all the sensitized crystals exhibit an initial period of fast decay followed by long, slow recombination. However, quantitatively the difference between the crystals has precluded a satisfactory analysis of the bleaching curves. On occasion, exponential bleaching has occurred similar to that observed in pure crystals, but with a much longer decay constant. However, the large majority do not follow the simple bleaching process. The third bleaching mechanism is of the bimolecular type. MacDonald[17] has analyzed some decay curves in the light of diffusion theory and has found that at room temperature $N=(A+Bt)^{-1}$, where N is the number of F centers and t=bleaching time. Unfortunately, all curves do not follow this bimolecular form of bleaching and the effect of crystal temperature and light intensity do not follow the predictions of the diffusion process.

Qualitatively, the difference between F centers in pure and sensitized salt follows from the nature of the holes and the complexity of the bleaching reaction, but quantitative explanations in both pure and sensitized salts require further work. Apparently, two types of hole configuration are required to account for the fast and slow bleaching of sensitized KBr in addition to possible recombinations between electrons and neutral bromine atoms because both bleaching rates are observed in crystals which have been previously irradiated in the F band to remove the holes. The recovery of the rapid bleaching rate during storage in the absence of F-band illumination requires that the configuration responsible for fast bleaching be capable of forming at room temperature in the absence of both ionizing and visible irradiation. Following the reasoning applied to plain crystals, it can be assumed that some of the hydrogen atoms will be located near or perhaps slightly attached to vacancies. These favorably located atoms readily trap electrons to reform the U centers and give rise to rapid bleaching. As these prospective U centers are used up, the bleaching rate slows down and depends on the rate with which hydrogen atoms and vacancies can come together. Interrupting the bleaching process with storage in the dark allows time for other hydrogens to reach favorable locations; consequently, rapid bleaching ensues when electrons are released by light. The limited number of prospective U-centers and the low rate with which they form would lead to an initial exponential type of decay. Unfortunately, many of the unexplained features observed with variations in light intensity, temperature, and color-center concentration result from unknown differences between crystals. Improved control is needed in the crystal preparation process for better quantitative results.

The authors wish to acknowledge the assistance of Miss Mary Denham who prepared many of the crystals and carried out the measurements, Lt. B. W. Downs who calculated the energy absorbed from Co^{60} gamma rays, Mr. Robert A. Levy who made the thermoluminescent measurements, and Professor F. Seitz who suggested the hydrogen-deuterium experiment.

[17] J. R. MacDonald, "Radiac detector DT-50 dosimeter reader and comparison scale" (unpublished report).

Structure in Optical Absorption of Barium Oxide Films*

R. J. ZOLLWEG
Cornell University, Ithaca, New York
(Received October 11, 1954)

Measurements of the optical absorption of BaO films at temperatures between 15°K and 370°C are reported. Four absorption peaks between 3.8 ev and 4.5 ev are found for measurements at liquid nitrogen temperature or below. At least the two strongest peaks are thought to be caused by exciton absorption. The sharpness of the structure is found to depend upon substrate temperature during evaporation and upon the temperature at which the optical measurements are made.

MEASUREMENTS of the optical absorption of BaO films were made between 3500A and 2000A at temperatures between 15°K and 370°C. The films were evaporated and the absorption measurements made in a liquid helium cryostat similar to that described by Duerig and Mador.[1] The evaporating unit was mounted inside the bottom of the liquid nitrogen shield, and the fused quartz substrate was held by a copper or steel holder which was connected to the bottom of the liquid helium container. Two $\frac{1}{4}$-inch apertures in the substrate holder exposed the quartz; the BaO film being evaporated onto the lower, with the upper serving as a "blank." The platinum or platinum plus 10 percent rhodium evaporating filament was outgassed for two hours at 1450°C and then coated with Mallinckrodt "ultra-pure" barium carbonate in a minimum of nitrocellulose, amyl acetate binder. This was heated slowly in the vacuum system and converted to the oxide with the substrate in the "up" position. The substrate was then lowered into the evaporating position and the BaO film evaporated at temperatures of 1225°C to 1300°C. The substrate was usually heated during the evaporation by a heater of the type used in receiving-type vacuum tubes. Liquid nitrogen was always kept in the cryostat liquid nitrogen container as a pump for water vapor when a BaO film was present. Under these conditions no change in transmittance, such as would occur by hydrolysis of the film, was observed though data from some films was taken over a period of several days.

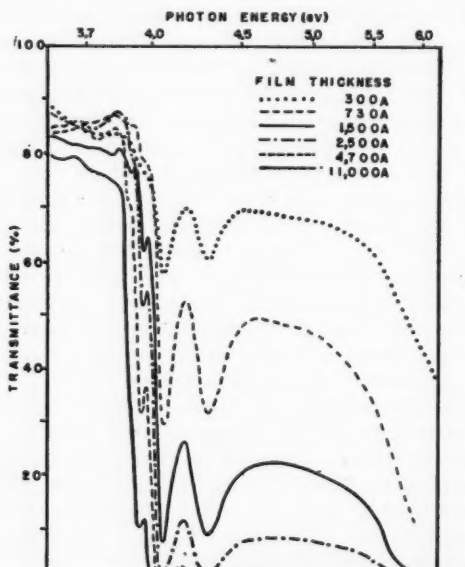


FIG. 1. Optical transmittance *vs* wavelength for BaO films of different thicknesses measured at -191 ± 4°C. Fused quartz substrate at 370 ± 10°C during evaporations.

The optical measurements were made using chopped light from a Nester Hydrogen arc, a 66-A/mm Bausch & Lomb grating monochromator, a 1P28 photomultiplier with a 33-cycle narrow-band amplifier and synchronous detector. A resolution of 6–8A was used and data were taken at intervals of 5 or 10A in the region of the observed structure.

When the optical transmittance was measured as a function of wavelength at liquid nitrogen temperature, much more structure was resolved than had been observed by Tyler[2] in the region near the absorption edge. This is shown in Fig. 1 for six BaO films of varying thickness. The relative thicknesses of these films were determined by time of evaporation and temperature of the evaporating filament using the data of Moore and Allison.[3] The absolute scale was selected by setting the absorption constant equal to 10^5 cm^{-1} at 2600A. This is the value obtained by Tyler, and weight measurements on some early films confirmed this magnitude of absorption constant. In addition to the strong absorption near 2000A, there also appear: (1) two strong absorption peaks at 2880A (4.30 ev) and 3050A (4.06 ev), (2) a small absorption peak at about 3140A (3.95 ev) on the edge of the 3050A absorption, and (3) a break in the slope of all of the curves for thicker

* This work was supported in part by the U. S. Office of Naval Research.
[1] W. H. Duerig and I. L. Mador, Rev. Sci. Instr. **23**, 421 (1952).
[2] W. W. Tyler, Phys. Rev. **76**, 1887 (1949). The general features of Tyler's work agree with those reported here. The observation of more detail here was the result of greater resolution attainable with the above system and further refinements such as heating the substrate during evaporation of the BaO and a lower measurement temperature.
[3] G. E. Moore and H. W. Allison, Phys. Rev. **77**, 246 (1950).

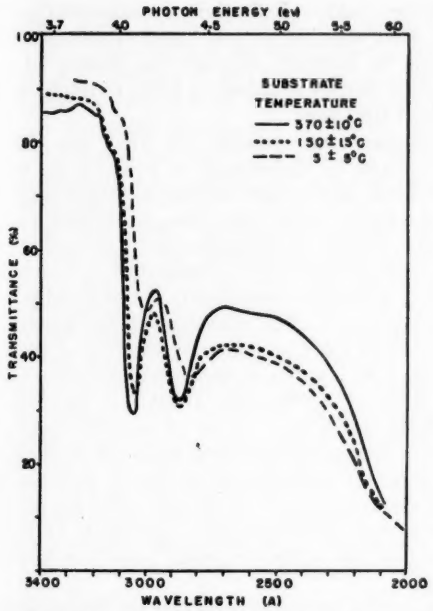


FIG. 2. Optical transmittance *vs* wavelength for different BaO films measured at -191 ± 4°C showing the effect of substrate temperature during the evaporation.

films at about 3200A (3.88 ev) suggesting the presence of an unresolved peak. The "wiggles" and overlapping of the curves for wavelengths greater than 3200A are probably caused by interference phenomena and in part by experimental error. Scattering is probably responsible for the reduction in transmittance in this region.

As shown in Fig. 2, the positions of the transmittance minima were observed to shift, and the relative transmittance at the two major peaks to change, as a function of the temperature of the fused quartz substrate during the evaporation of the BaO film. This is presumably associated with a change in crystal perfection in the films. Such a dependence of film perfection on substrate temperature has also been observed by Russell and Eisenstein[4] for evaporated BaO films on nickel substrates. There is some indication that raising the evaporating filament temperature tends to sharpen the structure. However, this may be through the mechanism of heating the substrate.

The optical absorption constant *vs* wavelength is plotted in Fig. 3 for three measurement temperatures. The liquid nitrogen temperature curve was determined by plotting the logarithm of the transmittance *vs* film thickness for each wavelength and determining the absorption constant by the slope of the best straight line through these points. For the other measurement temperatures, the curves were determined by assuming that all of the loss in transmittance was caused by absorption. From the experience with the liquid nitrogen temperature data it was found that little error would result from this treatment for sufficiently thick films. Transmittance measurements at 15 ± 5°K (using liquid helium as the coolant in the cryostat) showed structure only slightly sharper than at liquid nitrogen temperature. This presumably means that the sharpness of the absorption peaks at liquid nitrogen temperature is limited by zero-point lattice vibrations. It is possible, however, that lack of crystal perfection is responsible for a significant part of the observed width.

The two major absorption peaks are believed to be intrinsic to the BaO lattice. By use of Smakula's formula one calculates the concentration of absorption centers to be at least 1 percent and 0.5 percent for the peaks at 2880A and 3050A, respectively, after subtracting out a background "tail." Thus if this absorption were caused by defects or impurities, a very high concentration of these would be required. Defects associated with excess barium would perhaps be most likely. However, only about 0.04 percent excess barium is found in BaO single crystals colored by heating at 1120°C for 5 hours in a barium atmosphere.[5] In BaO films evaporated from platinum filaments, Moore and Allison[3] found less than 0.1 percent excess barium (their lower limit of detection). As a further experimental check on whether the structure could be associated with excess barium, a film was evaporated

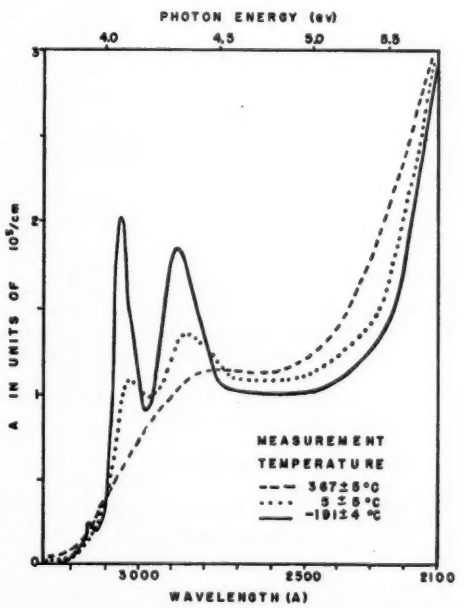


FIG. 3. Absorption constant (A) *vs* wavelength for BaO films measured at different temperatures. Fused quartz substrate at 365 ± 15°C during evaporations.

[4] P. N. Russell and A. S. Eisenstein, J. Appl. Phys. 25, 954 (1954).

[5] Sproull, Bever, and Libowitz, Phys. Rev. 92, 77 (1953).

from a filament consisting of a platinum plus 10 percent rhodium ribbon with a small tungsten ribbon spot-welded to it in a number of places. This arrangement was expected to give a film with the order of 10 percent excess barium.[6] The structure was not noticeably changed. An estimate of the minimum concentration of absorption centers associated with the peak at 3140A gives 0.03 percent while the value for the peak at 3200A would be considerably smaller.

The small widths of the major absorption peaks at half maximum (0.07 ev and 0.20 ev) and the fact that the structure sharpens with substrate temperature during evaporation of a film, appear to imply that the absorption is associated with the perfect lattice rather than with defects in the lattice. The sharpening of the structure at lower temperatures favors the view that these are exciton absorption peaks. Exciton absorption in BaO in the spectral region near 4.0 ev has been proposed to account for optical absorption and photoconductivity data,[7] photoemission,[8] and the temperature dependence of photoemission and photoconductivity.[9]

It should be noted that the absorption constant is at least as great as 300 cm^{-1} for single crystals of BaO prepared by a considerably different technique from these films,[5] and that the photoconductivity of these single crystals has a maximum in the region of the absorption peaks. Reflectivity measurements from single BaO crystals at low temperatures would be desirable to ascertain whether the density of these absorption centers is indeed as great in single crystals as it is in BaO films.

Overhauser[10] has pointed out that the observed structure can be understood in terms of an exciton model based on the tight binding approximation. The four absorption peaks arise from transitions between the ground state of the oxygen ion $(2p)^6$ and the four excited states of the $(2p)^5 3s$ configuration. The relative splittings of the four components can be fitted in terms of two constants: The exchange energy between $2p$ and $3s$ states and the spin-orbit splitting of the $(2p)^5$ ion core. A j-j coupling model is necessary. The absorption peaks at 3200A and 3140A are produced by transitions to excited states of angular momentum $J=0$ and 2, respectively. These peaks are weak because the transitions are forbidden by optical selection rules. The absorption peaks at 3050A and 2880A are produced by transitions to excited states for which $J=1$. They are allowed by optical selection rules and hence are strong.

The author is deeply indebted to Professor R. L. Sproull for his guidance and interest in this work and to Professors J. A. Krumhansl and A. W. Overhauser for discussions of interpretation.[11]

[6] Based on private communication from G. E. Moore.
[7] W. W. Tyler and R. L. Sproull, Phys. Rev. 83, 548 (1951).
[8] Apker, Taft, and Dickey, Phys. Rev. 84, 508 (1951).
[9] H. B. DeVore and J. W. Dewdney, Phys. Rev. 83, 805 (1951).
[10] A. W. Overhauser (private communication).
[11] Since this work was performed, Mr. Koji Okumura has kindly communicated his studies of optical absorption of BaO films. He has found results similar to the 5°C curve of Fig. 3.

PHYSICAL REVIEW VOLUME 97, NUMBER 2 JANUARY 15, 1955

Dynamics of Simple Lattices*

HERBERT B. ROSENSTOCK
U. S. Naval Research Laboratory, Washington, D. C.
(Received June 15, 1954)

The critical points (points where $\partial\omega/\partial x_i=0$ for all $i=1,2,3$) of the frequency ω of the lattice vibrations in wave number space (x_1,x_2,x_3), shown by Van Hove to exist for a very general class of crystals, are located for the monatomic simple, face-centered, and body-centered cubic lattices. The position and nature of the resulting singularities in the frequency distribution are found as a function of the ratio of force constants for second-nearest neighbors and nearest neighbors, and the qualitative features of the frequency distribution are thus determined. A method for using the information so obtained to determine the frequency distribution quantitatively is outlined.

I. INTRODUCTION

THE general problem we are concerned with is the finding of the energy eigenvalues of the Schrödinger equation for a system consisting of a very large number of coupled harmonic oscillators arranged in space in a simple lattice. Physically this system is interpreted as a crystal, the oscillators as the constituent atoms. Since the number N of atoms and hence the number of eigenvalues in a macroscopic crystal is of the order of Avogadro's number, an actual numerical evaluation of the latter is impractical even after an analytical expression for them has been obtained; and we are faced with the additional, usually more difficult, problem of finding the distribution function $g(E)dE$ of energy levels, defined as the density

* Preliminary results reported at the Washington Meeting of the American Physical Society, April 1954 [Phys. Rev. 95, 617 (1954)].

of energy levels to be found in the energy range $(E, E+dE)$. This function will determine the partition function $\int g(E)dE \exp(E/kT)$ and thus all other thermodynamic functions of the solid.[1] Among other fields in which frequency spectra are of interest may be mentioned optical properties (infrared absorption and Raman spectra), x-ray diffraction, superconductivity,[2] neutron scattering.[3]

Another way of formulating the problem, a semiclassical way, is to ask for the frequency eigenvalues ("normal modes") of the wave equation for the system described above considered classically, and for the "frequency distribution" $g(\omega^2)d\omega^2$, the fractional number of frequencies in the range $(\omega^2, \omega^2+d\omega^2)$, of them. The two formulations are equivalent mathematically, and also physically if Planck's assumption concerning the levels of an oscillator is made.[4] The semiclassical formulation is the one which is most used, even today, probably because the first paper on the subject[5] appeared before quantum mechanics was invented, and we shall generally follow this usage in this paper.

An exact calculation of the frequency spectrum has so far been possible for only very few crystals[6–14] (other than one-dimensional ones), and special, often physically unrealistic, assumptions had to be introduced. However, all frequency distributions found by exact calculation contained points of nonanalyticity: the two-dimensional crystals showed infinite logarithmic peaks, the three-dimensional ones infinite discontinuities in the slope. Van Hove[15] has traced the singularities to "critical points" (points where all $\partial\omega/\partial x_i=0$) of the surfaces of constant frequency in wave number space and has shown that a certain minimal number of critical points follows from the periodicity of ω in wave number space alone; hence that the frequency distribution of any crystal will show a minimum number of singularities.

In this paper we wish to show how the critical points, and from this the singularities in the frequency distribution, may be located for certain simple lattices. We use the square lattice first for an example in two dimensions and then treat the three-dimensional simple, face-centered, and body-centered cubic lattices. Finally we show how the information so obtained enables us to get a rather complete picture of the frequency distribution.

II. BACKGROUND AND NOTATION

We use the following abbreviations: CP for critical point(s), FD for frequency distribution, 1D, 2D, $\cdots i$D for one-dimensional, etc., sc, bcc, fcc for simple cubic, body-centered cubic, face-centered cubic.

The process of obtaining an expression for the frequencies of the normal modes of a monatomic lattice is straightforward and well known.[1] One sets the kinetic energy:

$$T=\tfrac{1}{2}m \sum_{\text{all } i,j,k} (\dot{u}_{ijk}^2+\dot{v}_{ijk}^2+\dot{w}_{ijk}^2),$$

and the potential energy:

$$V=[\tfrac{1}{2}\alpha \sum_{\substack{\text{nearest}\\ \text{neighbors}}} +\tfrac{1}{2}\beta \sum_{\substack{\text{2nd-nearest}\\ \text{neighbors}}}](d-d_0)^2.$$

Here $(u_{ijk},v_{ijk},w_{ijk})$ are the displacements in the (x,y,z) directions of the atom labeled according to its position by subscripts ijk. (If we are dealing with a 2D lattice the index k is not present.) d is the distance between two atoms and d_0 is the equilibrium distance; $d-d_0$ is to be expanded in the u, v, w and only the lowest nontrivial terms, leading to Hooke's law forces, are to be retained. m is the mass of each atom. α and β are force constants for nearest and second-nearest neighbors, respectively. One can then write down the Lagrangian equations of motion for each atom, solve these by standing waves with periodic boundary conditions, and thus obtain the secular equation:

$$\Delta(x_1,x_2,x_3;\alpha,\beta;\omega^2)=0, \qquad (1)$$

which determines the squared frequency ω^2. Strictly speaking, $x_i=p_i\pi/N^{\frac{1}{3}}$ is defined only for integer values of p_i between $-N^{\frac{1}{3}}$ and $N^{\frac{1}{3}}$, but since N is a very large number, x_i may be considered as a continuous variable running from $-\pi$ to π. Let us write down the determinant Δ explicitly for the cases of interest, calling the diagonal elements f_{ii} and the offdiagonal ones f_{ij}. For the 2D square lattice, we have[6]

$$f_{ii}=(1-c_i)+\sigma(1-c_1c_2)-\lambda,$$
$$f_{ij}=\sigma(1-c_i^2); \qquad (2)$$

[1] J. E. Mayer and M. G. Mayer, *Statistical Mechanics* (John Wiley and Sons, Inc., New York, 1940), Chap. XI.
[2] J. Bardeen, Phys. Rev. **80**, 567 (1950); H. Frohlich, Phys. Rev **79**, 845 (1950).
[3] G. Placzek and L. Van Hove, Phys. Rev. **93**, 1207 (1954).
[4] This equivalence follows essentially from the fact that semiclassical theory gives correct results when it deals with oscillators [see, e.g., L. Pauling and E. B. Wilson, *Quantum Mechanics* (McGraw-Hill Book Company, Inc., New York, 1935)], pp. 9, 30, 72. In connection with the problem at hand, this has been discussed explicitly by E. W. Montroll and D. C. Peaslee, J. Chem. Phys. **12**, 98 (1944), and it is brought up again only because conversations and the literature reveal that many persons seem to mistrust the results of the theory because it is usually formulated semiclassically. A remarkable example of this mistrust is the statement "*Die Schwingungen des idealen Gitters stellen ein quantenmechanisches System von makroskopischen Dimensionen dar; denn die Normalkoordinaten sind Amplituden von Wellen, die sich durch das ganze Gitter erstrecken. Das scheint eine bedenkliche Annahme.*" M. Born, *Festschrift Akademie der Wissenschaften, Gottingen* (Springer, Berlin, 1951), p. 1.
[5] M. Born and Th. von Kármán, Physik. Z. **13**, 297 (1912).
[6] E. W. Montroll, J. Chem. Phys. **15**, 575 (1947).
[7] M. Smollett, Proc. Phys. Soc. (London) **A65**, 109 (1952).
[8] W. A. Bowers and H. B. Rosenstock, J. Chem. Phys. **18**, 1056 (1950); **21**, 1607 (1953).
[9] G. F. Newell, J. Chem. Phys. **21**, 1877 (1953).
[10] C. M. Askey, thesis, University of North Carolina, Chapel Hill, North Carolina, 1951 (unpublished).
[11] J. P. Hobson and W. A. Nierenberg, Phys. Rev. **89**, 662 (1953).
[12] H. B. Rosenstock, J. Chem. Phys. **21**, 2064 (1953).
[13] H. B. Rosenstock and G. F. Newell, J. Chem. Phys. **21**, 1607 (1953).
[14] H. B. Rosenstock and H. M. Rosenstock, J. Chem. Phys. **21**, 1608 (1953).
[15] L. Van Hove, Phys. Rev. **89**, 1189 (1953).

for the 3D sc lattice, we have[14]

$$f_{ii}=(1-c_i)+\sigma(2-c_ic_j-c_ic_k)-\lambda,$$
$$f_{ij}=\sigma(1-c_j^2); \tag{3}$$

for the 3D bcc lattice, we have[16]

$$f_{ii}=1-c_1c_2c_3+\gamma(1-c_i^2)-\mu_B,$$
$$f_{ij}=(1-c_j^2)c_k; \tag{4}$$

for the 3D fcc lattice, we have[17]

$$f_{ii}=1-\tfrac{1}{2}c_i(c_j+c_k)+\sigma(1-c_i^2)-\mu_F$$
$$f_{ij}=\tfrac{1}{2}(1-c_j)^2. \tag{5}$$

Here $\sigma=\beta/\alpha$, $c_i=\cos x_i$, $\lambda=m\omega^2/2\alpha$, $\gamma=3\sigma/2$, $\mu_B=3\lambda/4$, $\mu_F=\lambda/2$ and the indices take on values 1, 2 in Eq. (2) and 1, 2, 3 in Eqs. (3), (4), and (5). Subscripts on μ_B and μ_F will be suppressed below.

We observe that in each case the secular equation (1) can be solved for λ or μ by elementary algebra, yielding three, or in the case of Eq. (2), two solutions ("branches") but that the expressions so obtained would be so complicated as to be of little use in finding the FD. Even the more restricted problem of finding the CP defined by $\partial\lambda/\partial x_i=0$ for all x_i is seen to present great algebraic difficulties when attacked by the straightforward approach of solving Eq. (1) for λ, differentiating with respect to x_1, x_2, x_3 and solving the resulting equations simultaneously. Henceforth, we shall therefore work mostly with the determinantal equation (1) itself rather than with the solutions which can in principle be explicitly obtained from it.

The first, and major problem of this paper, is to locate the CP of λ. Van Hove's paper,[15] in which the question of critical points and singularities in the frequency distribution caused by them was first generally treated, is based on a topological theorem due to Morse,[18] which van Hove states as follows:

> "Consider a function f defined on a closed topological manifold satisfying convenient conditions of differentiability and regularity. Assume f to be three times continuously differentiable and to have no degenerate critical points. Call index of a non-degenerate CP the number of positive eigenroots of the quadratic form in the Taylor expansion of f near the CP. Under these conditions the number of CP of index i is at least equal to the Betti number R_i of the manifold for the dimension i." By the Betti number R_i is meant "the maximum number of closed i-dimensional surfaces on the manifold which cannot be transformed into one another or into a point by a continuous deformation of the manifold." It is further stated that in two dimensions $R_0=1$, $R_1=2$, $R_2=1$, and in three dimensions $R_0=1$, $R_1=3$, $R_2=3$, $R_3=1$.

To the reader not versed in topology, it may not be clear from the above statement of the theorem, nor from the proof referred to,[18] that the predictions of the theorem for two- and three-dimensional spaces may also be obtained intuitively from extremely elementary considerations. Since in this paper more use is made of intuition than of mathematical rigor, a few words on this will be helpful.[19] Consider a function $f(x,y)$ periodic in both variables. (It helps to visualize this function as contour lines representing altitude, as on an ordinary geographical map.) Such a function is shown in Fig. 1. The unit cells of periodicity have without loss of generality been taken as squares and indicated in heavy lines. In the center cell, there will be at least one point where f takes on its maximum value[20]; we have called that point M and since in its neighborhood, the contours will be small closed curves, we have encircled it in Fig. 1a. From the required periodicity it follows that maxima will also appear at points labeled M_2, M_3, $\cdots$. Similarly, at least one minimum will exist in each cell, and these have been labeled m, m_1, m_2, $\cdots$ and also shown in Fig. 1a. Now consider the path of a traveler going from "mountain" M to "mountain" M_2 along some path A_1. He will have to start by going downhill, end up by going uphill, and reach a minimum somewhere in between, at a point which we may call (x_1,y_1). If he chooses an adjacent path A_2 instead, he will reach the minimum along that path at some adjacent point (x_2,y_2). Of all possible paths, there will exist one, A_i, at whose minimum the relation $f(x_i,y_i)>f(x_j,y_j)$ will be satisfied for all $j\neq i$. This point (x_i,y_i) is a saddle point and the contour is x-shaped there as indicated in Fig. 1b. In the same manner we see that another saddle point will appear between m and m_2. Other saddle points will appear at equivalent points in the other cells, by periodicity. In Fig. 1c these saddle points are shown connected up in the simplest possible way. We see that each cell has one maximum, one minimum and two saddle points; this is what Morse's theorem predicts for two dimensions. In three dimensions the surface $f=f_0$ through a saddle point is, near the saddle point, shaped like a double cone; adjacent surfaces are hyperboloids of one sheet for $f\lesseqgtr f_0$ and hyperboloids of two sheets for $f\gtreqless f_0$ and the saddle point is called an $S_\pm$ point accordingly. The reader who will construct (or visualize) the 3D model corresponding to Fig. 1 will have no difficulty convincing himself, by the same type of argument as we used for two dimensions, that each periodicity cell will con-

[16] P. C. Fine, Phys. Rev. **56**, 355 (1939).
[17] R. B. Leighton, Revs. Modern Phys. **20**, 165 (1948).
[18] M. Morse, Am. Math. Monthly **49**, 358 (1952); Trans. Am. Math. Soc. **27**, 345 (1925).
[19] A very similar "heuristic discussion" of Morse's theorem is given by E. W. Montroll, Am. Math. Monthly **61**, No. 7 (suppl), 46 (1954).
[20] The situation where the maximum value is taken on by a continuous family of points (a curve) can arise, but may be regarded as exceptional.

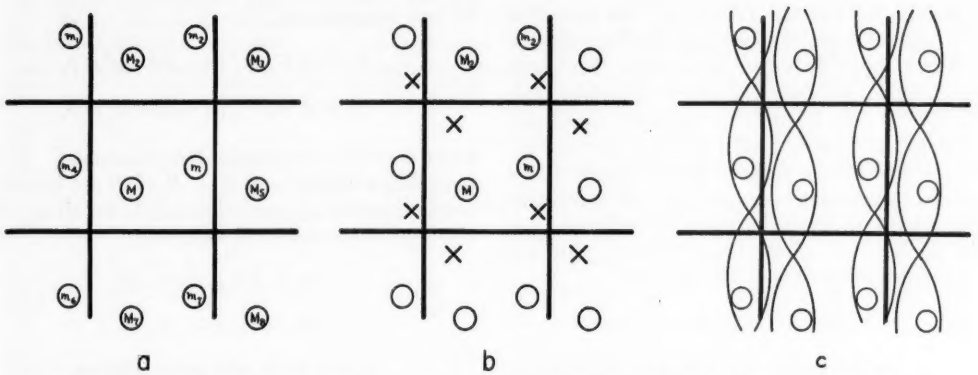


FIG. 1. Illustration of the predictions of Morse's theorem in two dimensions. Detailed explanation in text, Sec. II.

tain at least one maximum, one minimum, three S_+ points and three S_- points, in agreement with Morse's theorem.

Finally we wish to review briefly the conclusions of Van Hove concerning the nature of the singularities in the frequency distribution which each of the above CP will cause in a 3D crystal. This is done in Table I. Table I will enable us to sketch the frequency distribution as soon as we have located all the CP. Care must be exercised only in eliminating "generalized" CP which might give rise to less strong singularities.[21]

III. LOCATING THE CRITICAL POINTS

We have seen in the preceding section that we cannot hope to find the critical points by straightforwardly solving the three equations $\partial\lambda/\partial x_i=0$ simultaneously. In looking for another method we observe that λ as given by (1) with (2), (3), (4), or (5) depends on the x_i only through the $c_i=\cos x_i$. It follows first of all that we may restrict ourselves to considering the half-periodicity cube, $0\leq x_1, x_2 x_3\leq\pi$. Furthermore we may write $\partial\lambda/\partial x_i=(dc_i/dx_i)(\partial\lambda/\partial c_i)=-\sin x_i(\partial\lambda/\partial c_i)$. Since therefore $\partial\lambda/\partial x_i=0$ if and only if either $\sin x_i=0$ or[22] $\partial\lambda/\partial c_i=0$ it follows that we may have the following types of CP:

Type-1 CP All three $\sin x_i$ vanish.
Type-2 CP Two $\sin x_i$ vanish and one $\partial\lambda/\partial c_i$ vanishes.
Type-3 CP One $\sin x_i$ vanishes and two $\partial\lambda/\partial c_i$ vanish.
Type-4 CP All three $\partial\lambda/\partial c_i$ vanish.

Now $\sin x_i$ vanishes only when x_i is 0 or π. Since in the above classification no other mention appears of the x_i, we may henceforth consider the c_i as our variables and pay no more attention to the x_i. Henceforth when we speak of corners, faces, or other features of a cube, we shall mean the cube $-1\leq c_1, c_2, c_3\leq 1$. We may now say:

Type-1 CP appear at the 8 corners.
Type-2 CP appear on edges, at (1D) minima or maxima along the edges.
Type-3 CP appear on faces, at (2D) maxima, minima, or saddle points on the faces.
Type-4 CP appear inside the cube.

We have thus already located the minimum number 8 of CP predicted by Morse's theorem (the 8 Type-1 CP at the 8 corners). This of course is no accident but rather a consequence of the fact that Morse's theorem is arrived at entirely by periodicity considerations, and that the Type-1 CP are the ones that owe existence entirely to periodicity. The other type CP also depend to a greater or lesser extent on the detailed structure of the secular equation and are correspondingly harder to locate.

The locating of Type-2 CP is still rather easy. When two $\sin x_i$ vanish, four off-diagonal terms vanish, the determinant factors into the three diagonal terms which are linear in λ, and each of these can be considered individually. Each contains only one variable in which it is at most quadratic, the equation $\partial\lambda/\partial c_i=0$ is therefore easily solved, and the Type-2 CP have been found.

TABLE I. Nature of singularities in the frequency distribution (FD) of a three-dimensional (3D) crystal caused by ordinary (not generalized) critical points (CP). Here $\epsilon>0$; the table gives the first nonconstant term in the expansion of $g(\lambda)$ about λ_c; $\pm$ indicates the sign of the term's coefficient.

CP at λ_0	$g(\lambda_0-\epsilon)$	$g(\lambda_0+\epsilon)$
maximum M	$+\epsilon^{\frac{1}{2}}$	ϵ
minimum m	ϵ	$+\epsilon^{\frac{1}{2}}$
saddle point S_-	$-\epsilon^{\frac{1}{2}}$	ϵ
saddle point S_+	ϵ	$-\epsilon^{\frac{1}{2}}$

[21] "Generalized critical points" are defined in reference 15 as points near which the topological shape of the surfaces is the same as for ordinary CP but at which the equations $\partial\lambda/\partial x_i=0$ do not hold. The generalized critical point that appears in reference 15 at the absolute minimum of frequencies appears as an ordinary CP in our paper because we consider the distribution of $\lambda=m\omega^2/2\alpha$ rather than ω as does Van Hove. Since it is λ rather than ω that appears in the secular equation, this leads to other simplifications as well. Transition from the distribution $g(\lambda)d\lambda$ of λ to the distribution $G(\omega)d\omega$ of ω is made by noting that we must have $g(\lambda)d\lambda=G(\omega)d\omega$; hence $g(\lambda)[m/2\alpha]2\omega d\omega=G(\omega)d\omega$ or $G(\omega)=m\omega g(\lambda)/\alpha$.

[22] A possible exception to this statement might arise if $\partial\lambda/\partial c_i=\infty$ at a point at which $\sin x_i=0$. This does not occur.

The method for locating Type-3 CP, the ones that appear on a plane (square) is best explained by using the square lattice, Eqs. (1), (2), as an example. The equation is quadratic, so there will be two solutions; but, as explained in Sec. 2, we wish to avoid considering these in detail as their form will be rather complicated. However, their behavior on the edges is simpler and is easily found by equating c_1 or c_2 to 1 or -1; this causes the determinant to factor into two terms linear in λ. (If the Type-2 CP have been found explicitly, then this work has already been done.) We thus get:

$$\text{when } c_1=1,\begin{cases}\lambda_1=\sigma(1-c_2) & \text{from the } f_{11} \text{ term,}\\ \lambda_2=(1+\sigma)(1-c_2) & \text{from the } f_{22} \text{ term;}\end{cases}$$
$$\text{when } c_2=1,\begin{cases}\lambda_3=(1+\sigma)(1-c_1) & \text{from the } f_{11} \text{ term,}\\ \lambda_4=\sigma(1-c_1) & \text{from the } f_{22} \text{ term;}\end{cases} \tag{6}$$

and similar expressions when c_1 or c_2 are -1. In this way we determine the behavior of λ on the boundary of the square, and from this we shall attempt to find its behavior inside the square. One question which arises first however is whether at $(c_1,c_2)=(1,1)$ we shall join λ_1 to λ_3 and λ_2 to λ_4 or λ_1 to λ_4 and λ_2 to λ_3. (All four solutions take on the value of 0 at that point, so that at first sight either choice would seem possible. Points where two solutions take on the same value are called "contacts between branches" by Van Hove.[15]) In two dimensions this question can be easily and explicitly resolved. Consider the equation

$$\begin{vmatrix} A-\lambda & C \\ D & B-\lambda \end{vmatrix}=0. \tag{7}$$

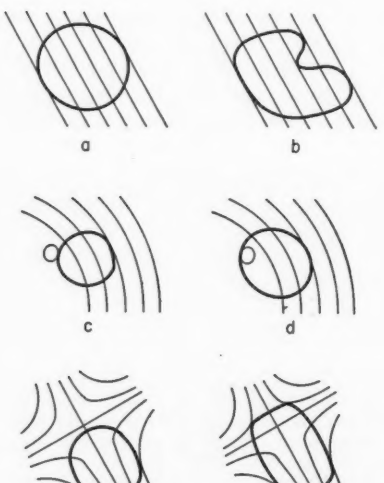


FIG. 2. Illustration of the relation between M_1 points on the boundary of a region and the two-dimensional critical points inside it. Detailed explanation in text, Sec. III.

Its two solutions are

$$\begin{aligned}2\lambda_+&=A+B+[(A-B)^2+4CD]^{\frac{1}{2}},\\ 2\lambda_-&=A+B-[(A-B)^2+4CD]^{\frac{1}{2}},\end{aligned} \tag{8}$$

where we wish to emphasize that, as usual, $[\ ]^{\frac{1}{2}}$ means the positive square root. If A, B, C, D are functions of two parameters x, y, note that $\lambda_+\geq\lambda_-$ for all x, y. When CD approaches 0, (8) approaches

$$\begin{aligned}2\lambda_+&=A+B+|A-B|,\\ 2\lambda_-&=A+B-|A-B|,\end{aligned}$$

or

$$\begin{aligned}\lambda_+&=A \text{ or } B, \text{ whichever is larger,}\\ \lambda_-&=A \text{ or } B, \text{ whichever is smaller.}\end{aligned} \tag{9}$$

This makes it clear that in (6) we must join λ_1 to λ_4 and λ_2 to λ_3. Observe that simply joining the solution arising from the f_{11} term for $c_1=1$ to the solution arising from the f_{11} term for $c_2=1$ would have been incorrect.

More generally we may assert: in order to properly identify solutions, it is sufficient to make sure that if $\lambda_1\leq\lambda_2\leq\lambda_3$ at any one point, then $\lambda_1\leq\lambda_2\leq\lambda_3$ at all points. For if, for any two points, there exists a path along which λ_1 is never equal λ_2, then λ_1 will be the larger at the end point if it started out larger at the initial point, and our rule for identification becomes necessary as well as sufficient; whereas if no such path connecting the two points exists, the two points must be separated by a continuous curve (or, in three dimensions, by a continuous surface) of points where the two solutions are equal; so that it then does not matter how the identification is made. Thus sufficiency is assured in either case. This argument and conclusion is equally valid in two or three dimensions; we shall have to keep it in mind also for locating Type-4 CP later.

We have now established the behavior of the solutions on the edges which bound the faces; in particular we know (from the work for Type-2 CP) the 1D maxima and minima along the bounding edges. We shall now show how this gives us considerable information concerning the 2D maxima, minima and saddle points in the square inside those edges. In Fig. 2, the heavy lines represent a closed curve bounding a region, and light lines, as before, are contours of constant frequency. (Although in most cases that concern us here, the region is a square, we have not drawn it so in order to emphasize the greater generality of the argument.) In Fig. 2a, no 2D CP is enclosed, and we find one 1D minimum and one 1D maximum along the boundary. This is the simplest possible case. In Fig. 2b, the boundary (or the contours or both) have been distorted and as a result four M_1 points now appear on the boundary (notation: and M_i point is an iD maximum or minimum). Note that at three M_1 points, the contours are tangent to the boundary from the outside (we call them $M_1{}^o$ points); and at one M_1 point, the contour is tangent to the

TABLE II. Possible 2D CP inside a region along whose boundary the 1D maxima and minima are known.

$M_1^o-M_1^i$	S_2-M_2	S_2	M_2	Illustration
−2	−2	0	2	Fig. 3a
		1	3	
		2	4	
		etc.		
0	−1	0	1	Fig. 3b
		1	2	
		2	3	
		etc.		
2	0	0	0	Fig. 3c
		1	1	Fig. 3d
		2	2	
		etc.		
4	1	1	0	Fig. 3e
		2	1	Fig. 3f
		3	2	
		etc.		
6	2	2	0	Fig. 3g
		3	1	
		4	2	
		etc.		
etc.				

boundary from the inside (we call this an M_1^i point). We can convince ourselves that by distorting the contours or the boundary, we can add M_1^i points and M_1^o points only in pairs, and that $M_1^o-M_1^i$ will remain equal to 2 under such distortions. Now assume that in a simple configuration as in Fig. 2a, a M_2 point is just outside the boundary (Fig. 2c) and observe what happens if the boundary is slightly distorted so that the M_2 point gets into the bounded region (Fig. 2d). We see that the result is the addition of one M_1^i point to the boundary and the removal of one M_1^o point from it. Similarly, Fig. 2e and 2f shows what happens to the boundary if it is distorted so as to enclose a new S_2 point (a two-dimensional saddle point); two M_1^o points get added to the boundary. We may summarize the results by writing

$$M_1^o-M_1^i=2+2S_2-2M_2. \tag{10}$$

This may be solved to give

$$S_2-M_2=\tfrac{1}{2}(M_1^o-M_1^i)-1. \tag{11}$$

From Eq. (11), Table II may be constructed. Table II tells what configurations of M_2 and S_2 points we may expect if we know the M_1^o and M_1^i, points on the boundary. (It should be added that in practice M_1^o and M_1^i points are readily distinguished by expanding the secular equation about them.) Some of the possibilities of Table II are illustrated in Fig. 3. For each value of $M_1^o-M_1^i$, an infinite number of possible configurations still remain, to be sure, but for the simple lattices we consider here and for comparatively low values of the force-constant ratio σ, only the simpler ones occur, and symmetry considerations enable us to decide which. Take, e.g., the square lattice again: λ as defined there is symmetrical about the line $c_1=c_2$, and it follows from this that if any S_2 or M_2 are contained in the square at all, at least one of them will appear on that line. A CP located on this diagonal is again easily found and investigated on account of the fact that the determinant simplifies along it. Detailed calculations are given in the next section.

The same general ideas that enables us to locate the Type-3 CP on the faces also apply to locating the Type-4 CP inside the cube. Although we have not been able to formalize our work by establishing a rule like (11), the study of two-dimensional M_2 or S_2 points on the surface of the cube, together with symmetry considerations, enable us, as is shown in the next section, to locate the fortunately small number of Type-4 CP that the lattices of interest show for fairly small σ.

IV. LOCATING THE CRITICAL POINTS—CALCULATIONS

A. Square Lattice

It is convenient to write

$$\begin{aligned}\bar{\mu}&=\lambda/(1+\sigma),\\ \tau&=\sigma/(1+\sigma),\\ p&=1-c_1,\\ q&=1-c_2,\end{aligned} \tag{12}$$

and to consider the square $0<p,\ q<2$ for $0<\tau<1$. The secular equation (1) with (2) beomes then

$$0=\Delta_2(p,q;\tau;\bar{\mu})$$

$$\equiv\begin{vmatrix} p+\tau q(1-p)-\bar{\mu} & \tau q(2-q) \\ \tau p(2-p) & q+\tau p(1-q)-\bar{\mu}\end{vmatrix}. \tag{13}$$

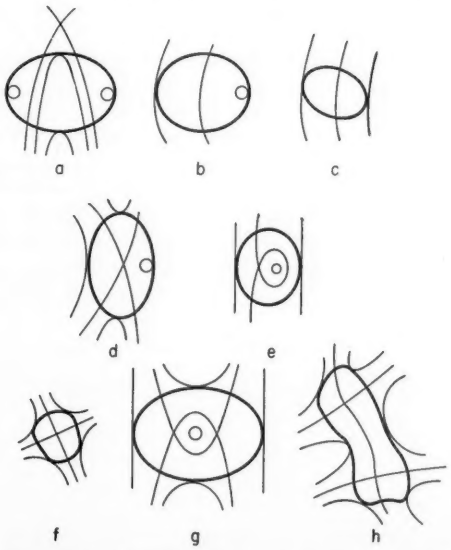


FIG. 3. Illustration of some of the configurations derived in Table II.

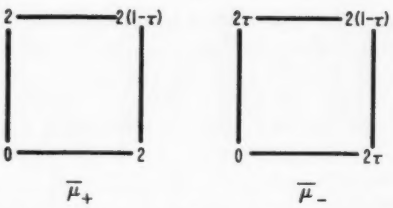


FIG. 4. Behavior of the solutions for the square lattice along the edges. Numbers show the values taken on by the two solutions at the corners.

Along the edge $p=0$ this reduces to

$$\bar{\mu}_1=\tau q, \quad \bar{\mu}_2=q; \tag{14a}$$

and along the edge $p=2$ this reduces to

$$\bar{\mu}_3=2-\tau q, \quad \bar{\mu}_4=2\tau+(1-2\tau)q, \tag{14b}$$

with similar expressions for the edges $q=0$, 2. Since the expressions (14) are linear, there will be no minima or maxima along the edges except possibly the end points (corners). Which solution is to be joined to which at the corners is determined by continuity at (0,2) and (2,0) and by the rule established in Sec. III at (0,0) and (2,2). Figure 4 shows the result. We see that $\bar{\mu}_+$ has along its boundary two minima at (0,0) and (2,2), and two maxima, at (0,2) and (2,0); $\bar{\mu}_-$ has along its boundary one minimum at (0,0) and one maximum at (2,2) if $\tau<\frac{1}{2}$, but has the same location of maxima and minima as $\bar{\mu}_+$ if $\tau>\frac{1}{2}$. To determine whether these are M_1^o or M_1^i points, we expand (13) about them and find: (0,0) is an M_1^o point in both $\bar{\mu}_+$ and $\bar{\mu}_-$ for all τ. (0,2) and (2,0) are M_1^o points for $\bar{\mu}_+$ for all τ, and M_1^o points for $\bar{\mu}_-$ for $\tau>\frac{1}{2}$. For $\bar{\mu}_+$, (2,2) is an M_1^i point for $\tau<\frac{1}{2}$ and an M_1^o point for $\tau>\frac{1}{2}$; for $\bar{\mu}_-$ it is an M_1^o point for $\tau<\frac{1}{2}$ and an M_1^i point for $\tau>\frac{1}{2}$. All this is illustrated in Fig. 5.

Next we look for 2D CP along the main diagonal (see the remark near the end of Sec. III) by setting $p=q$ and finding maxima or minima along that line by differentiation. We find only one, $\bar{\mu}=(1+3\tau)^2/8\tau$ located at $p=q=(1+3\tau)/4\tau$. This point is located inside the square for $\tau>\frac{1}{3}$, belongs to $\bar{\mu}_+$, and by expanding about it we find that it is a saddle point when $\tau>\frac{1}{3}$. Table II now tells us unambiguously that for $\bar{\mu}_-$, $(S_2,M_2)=(0,0)$ for all τ, and that for $\bar{\mu}_+$, $(S_2,M_2)=(0,0)$ for $\tau>\frac{1}{3}$. For $\bar{\mu}_+$ with $\tau>\frac{1}{3}$ there remain only the possibilities $(S_2,M_2)=(1,0)$, (3,2), (5,4), etc., since if there were an odd number of M_2, symmetry would demand that at least one of them show up on the main diagonal, and we reject all but (1,0) as too complicated to result from (13). The final sketches, Fig. 6, showing all CP, may now be made. We see that only three topological configurations occur; we shall call them L, M, N as defined in Table III.

B. Simple Cubic Lattice

It is convenient to write

$$\mu=\lambda/(1+2\sigma), \quad \rho=\sigma/(1+2\sigma),$$
$$p=1-c_1, \quad q=1-c_2, \quad r=1-c_3, \tag{15}$$

and to consider the cube $0<p,q,r<2$ for $0<\rho<1$. (1) with (3) then become

$$0=\begin{vmatrix} p+\rho(q+r)(1-p)-\mu & \rho q(2-q) & \rho r(2-r) \\ \rho p(2-p) & q+\rho(p+r)(1-q)-\mu & \rho r(2-r) \\ \rho p(2-p) & \rho q(2-q) & r+\rho(p+q)(1-r)-\mu \end{vmatrix}. \tag{16}$$

The behavior along the edges is established by equating two of the variables p, q, r to 0 or 2. Again we find that along each edge μ is linear in the third variable. From the symmetry of the determinant together with the rule for joining given in Sec. III, it follows that the cube surface of each of the three solutions of (16) will consist of two sets of three identical faces. The behavior on one set of these faces is found by setting one of the three variables equal to 0, on the other set by setting one of the variables equal to 2. Upon setting $r=0$, (16) gives

$$[\rho(p+q)-\mu]\Delta_2(p,q;\rho;\mu)=0, \tag{17}$$

TABLE III. Notation concerning configurations exhibited in Fig. 6 for the square lattice.

Range of τ	Configuration of $\bar{\mu}_+(\tau)$	Configuration of $\bar{\mu}_-(\tau)$
$0<\tau<\frac{1}{3}$	L	N
$\frac{1}{3}<\tau<\frac{1}{2}$	M	N
$\frac{1}{2}<\tau<1$	M	L

where Δ_2 is defined by (13) and analyzed in Sec. IVA. Upon setting $r=2$, (16) gives

$$[2-\rho(p+q)-\mu]\Delta_2(p,q;\sigma;\nu)=0, \tag{18}$$

with σ related to ρ by (15) and $\nu=(\mu-2\rho)/(1-2\rho)$. The cube faces are thus completely determined by the work of Sec. IVA and the simple expressions in the brackets of Eqs. (17), (18). After joining them up properly we find for the faces of the three solutions:

$$\mu_1(\rho)=\begin{cases}\rho(p+q) & \text{on front faces,}\\ (1-2\rho)\bar{\mu}_-(\sigma)+2\rho & \text{on back faces;}\end{cases}$$
$$\mu_2(\rho)=\begin{cases}\bar{\mu}_-(\rho) & \text{on front faces,}\\ (1-2\rho)\bar{\mu}_+(\sigma)+2\rho & \text{on back faces;}\end{cases} \tag{19}$$
$$\mu_3(\rho)=\begin{cases}\bar{\mu}_+(\rho) & \text{on front faces,}\\ 2-\rho(p+q) & \text{on back faces.}\end{cases}$$

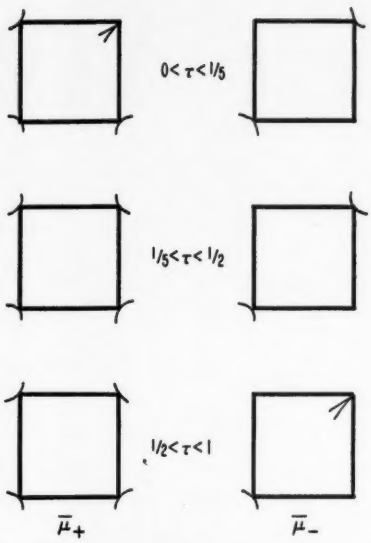


FIG. 5. M_1^i and M_1^o points for the square lattice.

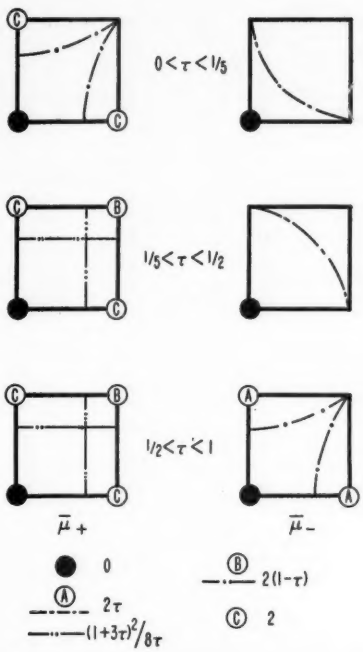


FIG. 6. Sketches of curves of constant frequency, square lattice.

Here $\bar{\mu}_\pm(\tau)$ are the functions sketched in Fig. 6; by "front" and "back" faces, we mean faces containing points $(p,q,r)=(0,0,0)$ and $(2,2,2)$, respectively. The two variables not constant on a face have for simplicity been written p, q. We may now use Table III to prepare an analogous table (Table IV) for the faces of the cube and also sketches of the surfaces (contained in Fig. 7). All Type-1 and Type-3 CP can be read off the sketches; there are no Type-2 CP. To locate the Type-4 CP we investigate the main diagonal of the cube and find one CP, $\mu_5=(1+6\rho)^2/16\rho$ located at $p=(1+6\rho)/8\rho$, and a double one, $\mu_6=1/4\rho$ located at $p=1/2\rho$. μ_5 is inside the cube if $\rho>1/10$, and by expanding we find that it is an S_3 point for $1/10<\rho<(2\sqrt{10}-5)/6\cong 0.221$ and an M_3 point otherwise. When it is an M_3 point the configuration becomes much more complicated and we shall therefore restrict ourselves to $\rho<0.221$ henceforth. This stipulation also eliminates all consideration of the CP μ_6, as it is inside the cube only when $\rho>\frac{1}{4}$. Study of the surfaces in Fig. 7 shows that the entry of an S_3 point into the cube at $\rho=1/10$ at the point $(2,2,2)$ is possible without change in the surface configuration only for the solution μ_3. Finally we remark that when $\rho=0$ we get two-dimensional continua[23] of critical points at the frequencies 0, 2. We are now able to exhibit in Fig. 7 sketches showing all critical surfaces for all ρ from 0 to 0.221.[24] The CP are listed in Table V.

TABLE IV. Configurations on curbe faces for simple cubic (sc) lattice, as exhibited in Fig. 7. Notation as in Table III.

Range of σ	Range of ρ	Configuration of $\mu_3(\rho)$ front face	$\mu_3(\rho)$ back face	$\mu_2(\rho)$ front face	$\mu_2(\rho)$ back face	$\mu_1(\rho)$ front face	$\mu_1(\rho)$ back face
$0<\sigma<\frac{1}{3}$	$0<\rho<\frac{1}{4}$	N	N	N	L	L	N
$\frac{1}{3}<\sigma<\frac{1}{2}$	$\frac{1}{4}<\rho<\frac{1}{3}$	N	N	N	M	L	N
$\frac{1}{3}<\sigma<\frac{1}{2}$	$\frac{1}{3}<\rho<\frac{1}{2}$	N	N	N	M	M	N

C. Body-Centered Cubic Lattice

This is more conveniently treated in terms of the original c_i rather than the p, q, r. One solution is found not to be linear along the edges, but no maximum or

TABLE V. Critical frequencies for the simple cubic lattice.

μ	CP	Range
0	m	all
$\mu_1=2\rho$	S_-	all
$\mu_2=4\rho$	S_+	all
$\mu_3=2-4\rho$	S_-	all
$\mu_4=2-2\rho$	S_+	all
$\mu_{\max}=2$	M	all
$\mu_5=(1+6\rho)^2/16\rho$	S_- $\frac{1}{10}<\rho<0.221$; M $0.221<\rho$	$\rho>\frac{1}{10}$
$\mu_7=(1+3\rho)^2/8\rho$	S_+	$\rho>\frac{1}{3}$
$\mu_8=2\rho+(1+\rho)^2/8\rho$	S_+	$\rho>\frac{1}{4}$

[23] This, and its effect on the frequency spectrum (see Sec. V) is well known. See, e.g., the discussion by Newell, reference 9. Newell has also predicted the entry into the cube of the S_3 point μ_5 when $\rho=1/10$.

[24] To the reader who has difficulty visualizing the connectedness of the surfaces from the rather congested drawing of Fig. 7, we suggest that he first prepare cubes (or half-cubes respectively) and drawn the appropriate lines on the surfaces; and that he then redraw the same figures on a sphere (or hemisphere respectively) or similar figure containing no sharp edges or corners to confuse him.

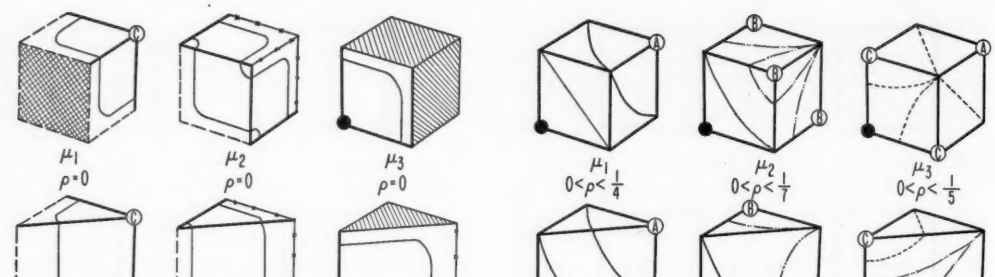


FIG. 7. Sketches of surfaces of constant frequency, simple cubic lattice.

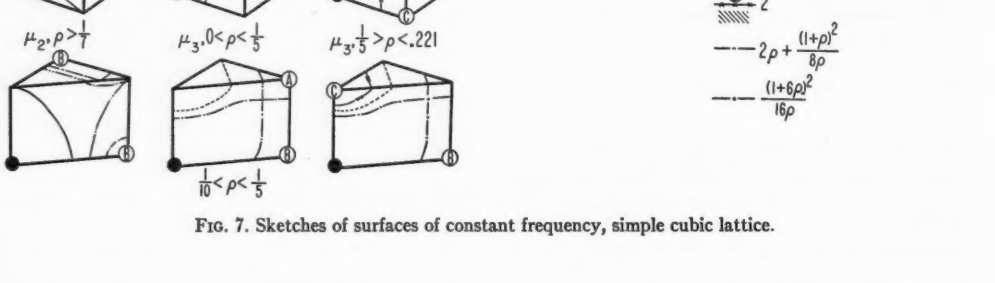


FIG. 8. Sketches of surfaces of constant frequency, body-centered cubic lattice.

minimum occurs for $-1<c_i<1$ unless $\gamma>\frac{1}{2}$; if we therefore take $\gamma<\frac{1}{2}$ there will be again no Type-2 CP. To find the behavior on the faces we again set $c_3=1$ for the front face and get

$$[1-c_1c_2-\lambda]\Delta_B(c_1,c_2;\gamma;\mu)=0, \tag{20}$$

Δ_B being what is left of the 2×2 determinant formed by f_{11}, f_{12}, f_{21}, f_{22}. For the back faces we set $c_1=-1$ and find

$$[1+c_1c_2-\lambda]\Delta_B(c_1,-c_2;\gamma;\mu)=0. \tag{21}$$

This shows that this time all six faces are alike except for a 90° rotation. If the two solutions arising from Δ_B in (20) or (21) are considered separately from the one arising from the bracket, it is found that the sufficient condition for joining (Sec. III) is satisfied and that in order to satisfy the convention that $\mu_1\leq\mu_2\leq\mu_3$ everywhere, the solution arising from the bracket should be labeled μ_2. The CP on the faces are found by the by now familiar method of first finding the nature of the M_1 points on the edges and then along the main face diagonal; each face is found to have one CP at $(c_i,c_j)=(0,0)$ as shown on the sketches on Fig. 8. Inside the cube along the main diagonal we find first of all critical points at $c_1=c_2=c_3=[-\gamma\pm(\gamma^2+18)^{\frac{1}{2}}]/9\cong\pm\sqrt{2}/3-\gamma/9$ where $\mu=2-2\gamma c_1-9\gamma c_1^2\cong1\pm0.63+7\gamma/9$, which upon expansion turn out to be saddle points for small σ. By symmetry, all diagonals are equivalent and hence the same saddle points appear on all of them. As we pass through the origin along the main diagonal, the first derivative does not vanish for any solution but the second derivative does; all solutions become equal there and in order to preserve the relations $\mu_1\leq\mu_2\leq\mu_3$ it is necessary to rearrange them properly. Even though the generalized CP resulting from this rearrangement will not cause a "strong" singularity (no discontinuity in the slope but only in the second derivative) in the FD, we must nonetheless study them carefully in order to determine the possible shapes of the surfaces throughout the cube. By expanding the original determinant about $(c_1,c_2,c_3)=(0,0,0)$, the solutions appear in the proper way directly. We thus find that μ_3 has a minimum, μ_1 a maximum, and μ_2 a double saddle point [the shape of the surface of μ_2 through (0,0,0) near (0,0,0) is a quadruple cone, the axis pointing to the points $(1,1,-1)(1,-1,1)$, $(-1,1,1)$, $(-1,-1,-1)$]. It turns out that these are all the critical points needed to give a consistent picture of all surfaces. They are shown in Fig. 8,[25] and listed in Table VI.

D. Face-Centered Cubic Lattice

The larger number of points- and line-contacts between branches makes the study of this lattice substantially more involved than of the preceding ones. Again we begin by investigating μ along the edges; then, upon considering the front faces $c_3=1$, we get

$$[2-(c_1+c_2)-\mu]\Delta_F(c_1,c_2;\sigma;\mu)=0, \tag{22}$$

with Δ_F defined in analogy with (20). At the back faces $c_3=1$, we get

$$[2+(c_1+c_2)-\mu]\Delta_F(-c_1,-c_2;\sigma,\mu)=0, \tag{23}$$

and so again only one face need be studied in detail. The one CP on the main diagonal of the face, at $c_1=c_2=(5+4\sigma)/4(1+\sigma)$ where $\mu=(5+4\sigma)^2/16(1+\sigma)$, is found to be an S_2 point for $\sigma<\sigma^+\cong0.149$ and an M_2 point for $\sigma>\sigma^+$. [Precisely, σ^+ is defined as the solution of the equation $\sigma(5+4\sigma)(3+4\sigma)=3$.] The generalized CP $(3+4\sigma)/(2+2\sigma)$ and $\frac{3}{2}$ result from interchanging solutions along the line contact running from $(-1,0)$ to $(0,-1)$ on each face. The configuration on the surface changes when at $\sigma=0$ these two frequencies become equal and again when at $\sigma=(-2+\sqrt{3})/4\cong-0.067$ when $(5+4\sigma)^2/16(1+\sigma)$ and $\frac{3}{2}$ become equal. These changes do not seem to affect the structure of the surfaces inside the cube qualitatively, however. After ascertaining the behavior of the solutions at the corners and other M_1 points on the edges, the configurations

TABLE VI. Critical frequencies for the body-centered cubic (bcc) lattice. Good to $\gamma=\frac{1}{4}$. When $\gamma=0$, 0 and γ coalesce, and 2 and $2+\gamma$ coalesce to give a continuous line of CP.

μ	CP
0	m
γ	S_-
$\mu_-\cong0.37+7\gamma/9$	S_+
1	S_-
$1+\gamma$	(S,M,m)(generalized)
$\mu_+\cong1.63+7\gamma/9$	S_-
2	S_+,M,M
$2+\gamma$	M

TABLE VII. Critical frequencies for the face-centered cubic (fcc) lattice. Only ordinary (not generalized) CP are shown.

σ	μ	CP
$-\frac{1}{4}<\sigma<-0.067$	0	m
	$\frac{1}{2}+\sigma$	S_-
	$1+\sigma$	S_+
	1	M,S_-
	$(5+4\sigma)^2/16(1+\sigma)$	M
	$\mu^*\cong1.8$	S_+
	$2+\sigma$	M
	2	M
$-0.067<\sigma<0$	$(5+4\sigma)^2/16(1+\sigma)$	S_+
	otherwise unchanged	
$\sigma=0$	1 and $1+\sigma$ coalesce to give a continuous line CP	
	2 and $2+\sigma$ coalesce	
$0<\sigma<0.149$	0	m
	$\frac{1}{2}+\sigma$	S_-
	1	S_-,S_+
	$1+\sigma$	M
	$(5+4\sigma)^2/16(1+\sigma)$	S_+
	$\mu^*\cong1.8$	S_+
	2	M
	$2+\sigma$	M

[25] The surfaces are seen to be more complicated than in Fig. 7. The essentials of μ_2 have been described in the text. μ_1 may be visualized best by paying most attention to the surface $\mu=0.37+7\sigma/9$. It consists of one closed surface which encloses the center of the cube and which touches, at four points located on the diagonals, four cup-shaped surfaces each of which surrounds one corner of the cube. μ_3 has a similar structure.

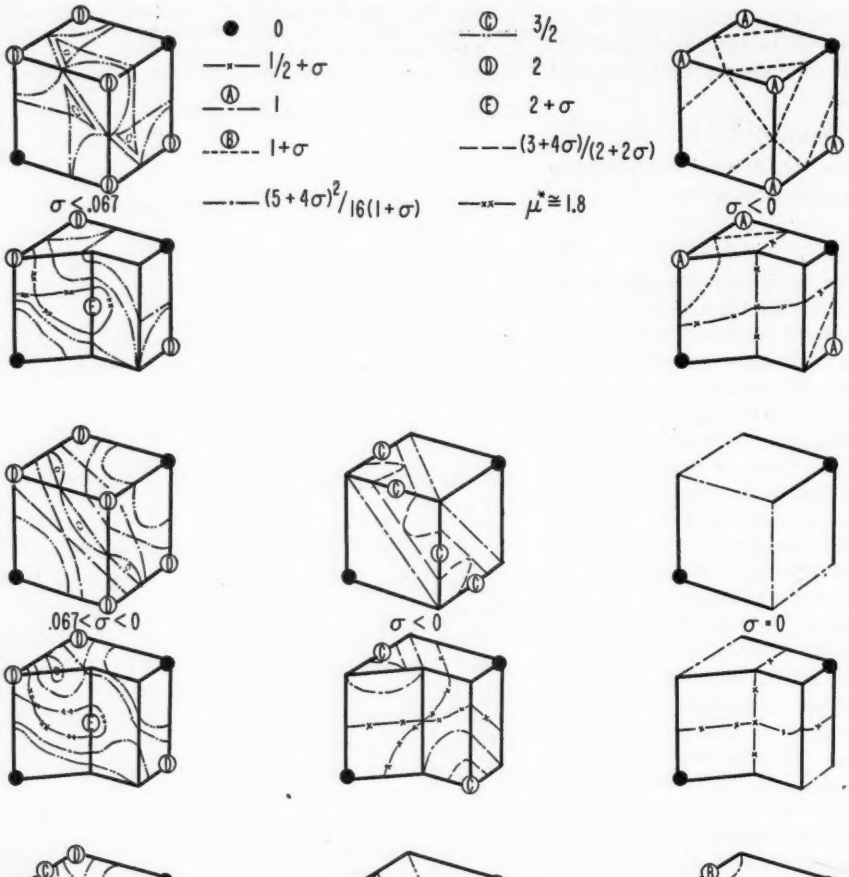


FIG. 9. Sketches of surfaces of constant frequency, face-centered cubic lattice.

shown on the surfaces of the drawings of Fig. 9 are obtained for $\sigma<\sigma^+$. Since in contrast to the other lattices considered in this paper, the fcc is not unstable for all $\sigma<0$, we include sketches of solutions for $\sigma<0$ whenever they are different from the ones for $\sigma>0$. Along the main diagonal inside the cube, each solution is found to have a CP at the origin. In two cases $\mu=\frac{1}{2}+\sigma$ and the CP is a S_3 for $\sigma<\frac{1}{6}$, in the third case $\mu=2+\sigma$ and the CP is a M_3 for all σ. This gives a consistent interpretation for the solutions μ_2 and μ_3, but the fact, revealed by expanding, that μ_1 is smaller than 2 in the neighborhood inside the cube of the points $(1, 1, -1)$ indicates that another S_3 point must exist inside the cube. This point is not located on the diagonal that runs from $(0,0,0)$ to $(1, 1, -1)$ and so we have not been able to locate it or find the exact frequency μ^* of the critical

surface. However for $\sigma=0$ the minimum value of μ_1 along the diagonal $c_1=c_2=-c_3$ is $\mu_0=6(16+3\sqrt{2})/17 \cong 1.79$ and this value is not very sensitive to changes in σ; we may therefore conclude that $1.7<\mu^*<(2$ or $2+\sigma$, whichever is smaller). The final results are shown[26] in Fig. 9 and Table VII.

V. THE FREQUENCY DISTRIBUTION

The FD for the sc, bcc, and fcc lattices can be sketched directly from the information given in Tables V, VI, VII with the help of Table I. The sketches are given in Figs. 10, 11, and 12 respectively. The square lattice which has been rather fully treated by Montroll[6] has been included in this paper primarily to illustrate the problems in a simple way and will not be treated further. It is interesting to note in Fig. 10 that all new singularities that appear as ρ increases "grow out of" old singularities; this may be seen from Fig. 7 to be due to the fact that new CP invariably enter the cube by way of, or form out of, old CP. When we get a continuous surface of CP for the sc lattice (Fig. 7) and an inverse-square-root singularity, which is the singularity characteristic of 1D problems, appears in the FD. Physically this is due to the fact that, with nearest-neighbor interaction only, the motion in the x-direction of an atom in a sc lattice is independent of all atoms other than those in the x-direction, so that we are in effect dealing with a 1D problem. Similarly when $\sigma=0$ we get a continuous line of CP for the bcc and fcc lattice (Figs. 8 and 9, respectively) and a logarithmic singularity, which is the singularity characteristic of 2D problems, appears in the FD. For this we know no explanation as simple as for the inverse square root singularity in the sc lattice.

The only previous work that we can compare our results with nontrivially is Leighton's[16] numerical-mechanical computation of the fcc FD for two values of σ. Figure 13 shows Leighton's results for $\sigma=-0.1$ (replotted so that the abscissa is μ rather than $\mu^{\frac{1}{2}}$) together with the singularities found in this paper. Agreement is good; some of the singularities actually show up as discontinuities in the slope, though not as infinite ones (1,1.47), and others (0.4,0.9,λ^*,1.8) as rather sharp, though not discontinuous changes in the slope; some are slightly displaced. Clearly such smoothing would be expected to result from any numerical treatment. The agreement for $\sigma=0$ (not shown in Fig. 13) is equally good. The peak at $\mu=1$ is definitely there, though of course not infinitely high.

We cannot compare our work of that of Fine[26] since his choice of $\gamma=1$ falls outside our range. Our results

[26] The structure of μ_3 and μ_2 is comparatively easy to visualize. The essential surface is the one that goes through the S_3 point (0,0,0). The change in the surface of μ_3 as σ goes through 0 results from the fact that $1+\sigma \gtrless 1$ as $\sigma \gtrless 0$. The structure of μ_1 is somewhat similar to μ_1 and μ_3 in the bcc lattice (see reference 25): The surface $\mu=\mu^*$ consists of a closed surface which encloses the origin and to which six cup-shaped appendages are attached.

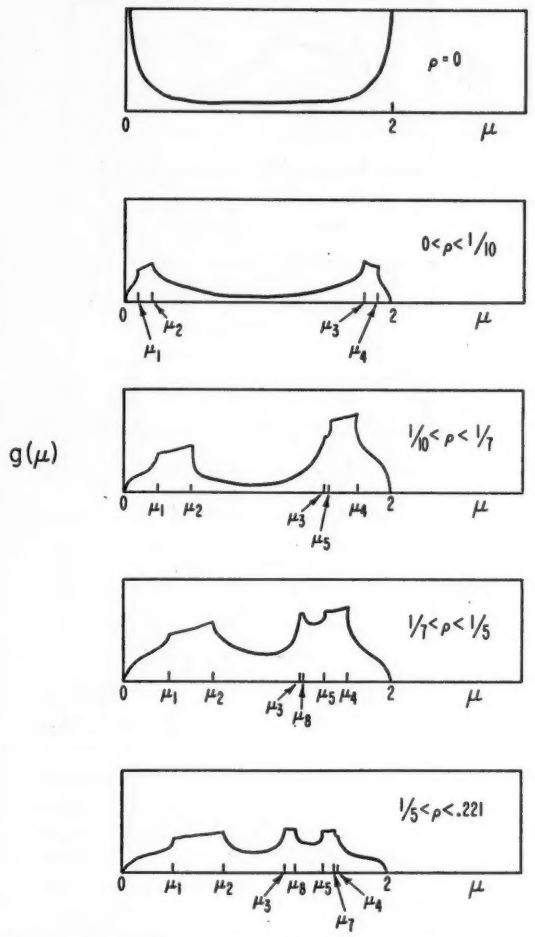


FIG. 10. Sketches of frequency distribution, simple cubic lattice.

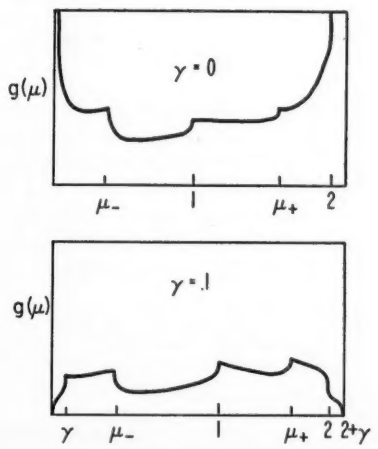


FIG. 11. Sketches of frequency distribution, body-centered cubic lattice.

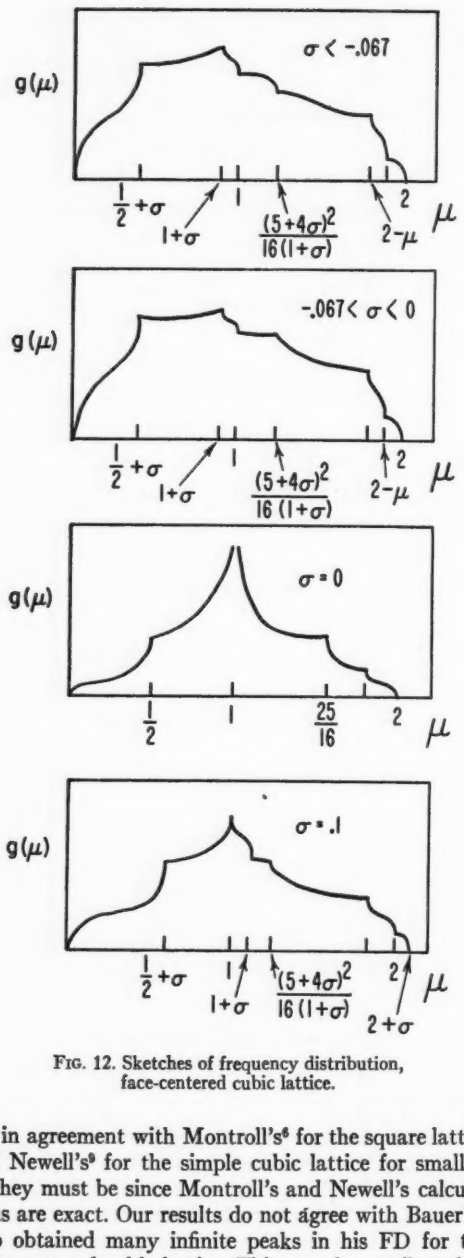


FIG. 12. Sketches of frequency distribution, face-centered cubic lattice.

are in agreement with Montroll's[6] for the square lattice and Newell's[9] for the simple cubic lattice for small ρ, as they must be since Montroll's and Newell's calculations are exact. Our results do not agree with Bauer's[27] who obtained many infinite peaks in his FD for the body-centered cubic lattice. This may be attributed to Bauer's use of Houston's[28] approximation.

[27] E. Bauer, Phys. Rev. **92**, 58 (1953).

[28] W. V. Houston, Revs. Modern Phys. **20**, 161 (1948). The idea on which the approximation is based is the replacement of the integration over the cube $0<x_1, x_2, x_3<\pi$ in the exact expression

$$g(\omega^2)=\int\int\int\delta[\omega^2-F(x_1x_2x_3)]dx_1dx_2dx_3$$
$$=\int\int\int\delta(\omega^2-F)r^2dr\,\sin\theta d\theta d\phi,$$

where $\omega^2=F(x_1x_2x_3)$ is the expression for ω^2, by a sum $\sum_n c_n\int\delta(\omega^2-F)r^2dr$ over a finite number of directions (ϕ_n,θ_n), and interpolating and normalizing properly. What makes the approximation useful is the fact that for certain directions (ϕ_n,θ_n) the expression $\omega^2=F$ is simple enough for the r-integration to be carried out, whereas in general it is not. The concomitant difficulty is that these directions are the ones running from the origin to a corner; i.e., from and to a CP. As a result, the triple integration over the CP is in effect replaced by a single one, and the singularities that appear in the 3D FD are of the nature that is correctly obtained in the 1D case, that is to say, infinite peaks of inverse square root nature. The appearance of spurious infinite peaks in Houston's approximate method has been previously discussed in connection with two specific 2D problems in reference 8 and by T. Nakamura, Progr. Theoret. Phys. (Japan) **5**, 213 (1950). It should be said that Bauer seems to be aware of the shortcomings of the method.

VI. THE ABSOLUTE VALUE OF THE FD

In the foregoing we have, for certain lattices and force constant ratios, found the frequencies at which the FD has singularities and established the nature of each singularity. We have not, however, done anything to find the absolute value of the FD at any point, except trivially at the end points where $g=0$. In this section we indicate briefly an approximate method by which, once the position and nature of the singularities is known, the FD may be found quantitatively with what we believe to be little labor for the accuracy attained. One sample calculation which bears out this belief has been performed.

The idea is to modify Montroll's approximate method, which was discovered at a time when the existence of singularities in the FD was not known, in such a fashion as to take proper account of the singularities. This occurred simultaneously to Lax and Lebowitz[29] and the writer.[30] Montroll's original method[31] may be described thus: Expand the frequency distribution,

$$g(\lambda)=\sum_n p_n\phi_n(\lambda), \tag{24}$$

the $\phi_n(\lambda)$ being known functions, such as Legendre polynomials, powers of λ, etc.; the p_n are coefficients undetermined at this point. Multiply each side by λ^k and integrate over λ; the result is

$$m_k=\sum_n a_{kn}p_n, \tag{25}$$

where $m_k\equiv\int\lambda^k g(\lambda)d\lambda$, called the kth moment of the FD, is seen to be the average value of λ^k, and $a_{kn}=\int\lambda^k\phi_n(\lambda)d\lambda$ can be directly computed. If we knew the first s moments, we could cut the sum off at $n=s$ and solve the set of s equations (25) for the expansion coefficients p_n. Equation (24) would then give us the FD approximated by a linear combination of a finite number of the functions $\phi_n(\lambda)$. The point is that the average value of λ^k, which may also be written

$$m_k=\int\int\int\tfrac{1}{3}\sum_{i=1}^{3}\lambda_i^k dx_1dx_2dx_3, \tag{26}$$

[29] M. Lax and J. L. Lebowitz, Phys. Rev. **95**, 617(A) (1954); Phys. Rev. **96**, 594 (1954). Detailed calculations for one 2D crystal are given.

[30] See footnote to title.

[31] E. W. Montroll, J. Chem. Phys. **10**, 219 (1942).

may indeed be easily computed even when λ_i is not known explicitly; for (1) may be written as $M\psi=\lambda\psi$, where M is the matrix of the determinant Δ with λ set equal to 0; multiplication from the left by M^{k-1} gives $M^k\psi=\lambda^k\psi$. An elementary theorem of matrix algebra then states that

$$\mathrm{Tr}(M^k)=\sum_{i=1}^{3}\lambda_i^k;$$

hence from (26)

$$m_k=\int\int\int \tfrac{1}{3}\,\mathrm{Tr}(M^k)dx_1dx_2dx_3. \qquad (27)$$

To take account of the singularities we merely replace (24) by something like

$$g(\lambda)=\sum_{j=1}^{m}F_j(\lambda)1(\lambda_{j-1},\lambda_j), \qquad (28)$$

where

$$1(a,b)=\begin{cases}1 & a<\lambda<b\\ 0 & \text{otherwise,}\end{cases}$$

and $\lambda_0=0<\lambda_1<\lambda_2<\cdots<\lambda_m=\lambda_{\max}$ are the values of λ at which singularities in the FD occur, and F_j is a function which contains a number of undetermined coefficients and which has the correct behavior at λ_{j-1} and λ_j. The coefficients of the first term $|\lambda-\lambda_j|^{\frac{1}{2}}$ near any singularity can in principle be found by expansion of the determinant about the CP that causes the singularity; in practice this is easiest for Type-4 CP and most difficult for Type-1 CP on account of degeneracy; it may be easier to keep that coefficient undetermined.

The example we have treated is the sc lattice for $\rho=0.05$. The location and nature of singularities is found from Tables V and I. We assume the following f_j:

$$f_1=c_0+c_1\mu,$$
$$f_2=c_2+c_3\mu,$$
$$f_3=c_4(\lambda-0.1)^{\frac{1}{2}}+c_5(0.9-\lambda)^{\frac{1}{2}}+c_6,$$
$$f_4=c_7+c_8\mu,$$
$$f_5=c_9+c_{10}\mu.$$

That is to say, we took the additional shortcut of assuming that the FD was a straight line in the ranges (0,0.05) and (0.95,1) and ignoring the square-root-type singularities which appear at the end points of those ranges and which, according to the recipe given in the preceding paragraph, we should have taken into account. In the case at hand this is justified by the fact that the shortness of the range assures that these singularities will not affect the shape of the FD much. The total number of undetermined coefficients included was thus eleven; one of those, c_4, could be easily found from the determinant, and six more were eliminated by the use of the conditions that $g(0)=g(1)=0$ and that g is continuous for all μ; four simultaneous equations of the form (25) had therefore to be solved. The moments were taken from Montroll.[21] μ has been normalized to run from 0 to 1. The results are shown in Fig. 14, the singularities at 0, 0.05, 0.95, 1 having been restored. Newell's[9] exact FD is also shown. In view of the fair agreement obtained by the use of only four free parameters, it may be hoped that more complicated approximating functions, together with the use of a high-speed computer to solve the larger number of simultaneous equations, will give accurate results. We hope to present results of such calculations in the future.

I wish to thank Professor W. A. Bowers for discussions and criticism and Mrs. H. M. Rosenstock and Mr. I. H. Blifford for help with computations.

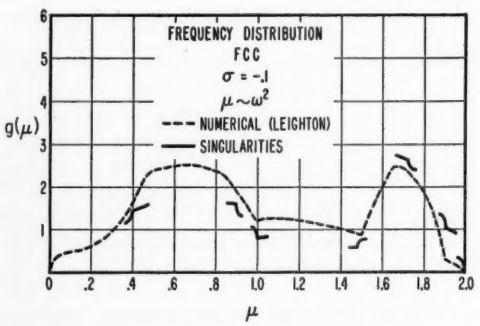


FIG. 13. Comparison between Leighton's numerical work and the results obtainable from this work for the square lattice.

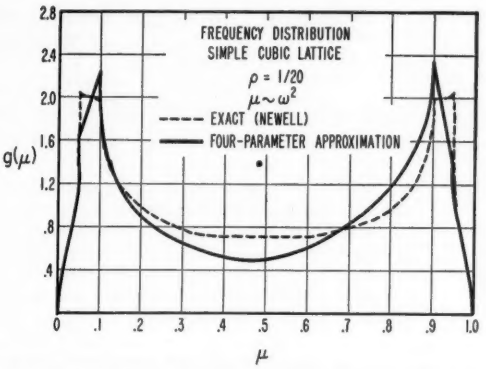


FIG. 14. Frequency distribution obtainable by applying Montroll's approximate method to the results of this work, compared with Newell's exact result for the simple cubic lattice, small ρ.

Neutron Diffraction Studies of the Magnetic Structure of Alloys of Transition Elements

C. G. SHULL AND M. K. WILKINSON
Oak Ridge National Laboratory, Oak Ridge, Tennessee
(Received September 7, 1954)

Neutron diffraction results are presented for a series of alloys of transition elements as an investigation of their magnetic structure. Ferromagnetic disorder scattering has been found for alloys in the disordered state and magnetic superstructure scattering is observed for ordered alloys. Such magnetic scattering when combined with magnetization data offers information on the individual atomic magnetic moments present in the alloy. In general the atomic magnetic moments deviate from their pure elemental values as a function of alloy composition. Scattering data and magnetic moment information are given for members of the Fe–Cr, Ni–Fe, Co–Cr, and Ni–Mn series of alloys.

INTRODUCTION

THE electronic structure existing in alloys of the transition elements has been the subject of considerable discussion, not always without controversy, in the literature of the last decade. Various viewpoints of the structure of the transition elements themselves have been presented and their proponents have in many cases extended the different treatments into the alloy field. Among the more recent review articles of work in this field are those of Goldman[1] and Hume-Rothery and Coles[2] wherein extensive references to earlier research publications will be found.

Experimentation in this field has centered primarily about magnetic measurements, since for many atomic systems, this type of information can lead to an unambiguous picture of the electronic structure. Such measurements are, however, invariably representative of the macroscopic average over the whole component system and offer no information whatsoever about details from one atom to another. In principal, neutron scattering can distinguish between the different electronic configurations of different atoms via a magnetic interaction and thus offer some guidance in the selection or rejection of particular theoretical viewpoints.

A number of different alloys systems have been studied in this fashion over the past several years, and although the studies are by no means complete since other equally interesting systems remain for investigation, it was thought desirable to summarize the present results. The alloys to be discussed in the present report are those shown superimposed on the Slater-Pauling curve of Fig. 1. It was noticed some fifteen years ago[3] that the ferromagnetic moments of the pure elements and many alloys tended to follow such a general curve with respect to average outer electron concentration. Some exceptions were noted, however, and typical of these is the Co–Cr system which possesses magnetic moments that vary as shown in the figure. In addition to the Co–Cr, Fe–Cr, and Ni–Fe systems, data have been obtained on two alloys of definite composition, Ni_3Fe and Ni_3Mn. These two alloys are known to form ordered superstructures, and the investigation gives information concerning the effect of atomic ordering on the magnetic structure.

NEUTRON SCATTERING CONSIDERATIONS

It is well known that elements of the transition group form alloys possessing solid solubility over extended regions of composition. Such alloys exist usually in the *disordered* form in which the solute atoms reside at random on solvent atom sites. For certain compositions, however, the alloy may take on an *ordered* form wherein the atoms of one species become located at specific crystallographic positions. An arrangement of the latter type is usually brought about by suitable thermal treatment of the alloy, and its existence can be detected by the appearance of superstructure line intensity in the x-ray or neutron diffraction pattern. Both disordered and ordered alloys are included in the present study and it will be desirable to give expressions for the scattered intensity or cross section appropriate to the two cases.

Ferromagnetic Disorder Scattering

When neutrons are scattered from a ferromagnetic alloy of disordered form, there should exist a ferromagnetic disorder scattering whose cross section or intensity will depend upon the difference in the magnetic scattering amplitudes of the two species of atoms. If the incident neutron beam is unpolarized, it can be shown that this cross section for an unmagnetized sample in which the ferromagnetic domains are oriented at random is given by

$$d\sigma = \tfrac{2}{3}n(1-n)(p_1-p_2)^2, \qquad (1)$$

wherein n and $(1-n)$ are the fractional abundances of the two species and p_1 and p_2 are the magnetic scattering amplitudes. The factor $\frac{2}{3}$ arises because of the randomness of the magnetization vector relative to the scattering vector for the unmagnetized sample. Furthermore

[1] J. E. Goldman, Revs. Modern Phys. 25, 108 (1953).
[2] W. Hume-Rothery and B. R. Coles, Advances in Physics 3, 149 (1954).
[3] J. C. Slater, J. Appl. Phys. 8, 385 (1937); L. Pauling, Phys. Rev. 54, 899 (1938).

the magnetic scattering amplitude will depend upon the atomic magnetic moment μ according to

$$p=\tfrac{1}{2}(e^2/mc^2)\gamma\mu f$$
$$=0.270\times10^{-12}\mu f, \quad (2)$$

where γ is the neutron magnetic moment and f is the magnetic form factor descriptive of the spatial origin of the atomic magnetic moment. Thus it is seen that a determination of the ferromagnetic disorder scattering can yield the difference between the magnetic moments of the two types of atoms present in the alloy. If at the same time the average magnetic moment of all the atoms is known from magnetic saturation data, *viz.*,

$$\bar{\mu}=n\mu_1+(1-n)\mu_2, \quad (3)$$

then expressions (1) and (3) may be combined to permit solution of μ_1 and μ_2 within an algebraic sign ambiguity. Unfortunately the intensity of scattering represented by Eq. (1) does not give information on the sign of the moment difference; hence alternative sets of solutions for the individual moments are obtained. There do exist experimental techniques for removing this ambiguity which involve the use of polarized neutron beams or a study of the paramagnetic scattering at high temperatures, but as yet these experiments have not been performed. Throughout the data presentation, both sets will be given, and it is necessary to use independent arguments in assigning the correct solution.

The ferromagnetic disorder scattering will appear in the diffuse background of the neutron diffraction pattern and its extraction from the various other diffuse scattering components such as multiple scattering, thermal disorder scattering, and isotopic or nuclear spin disorder scattering offers considerable difficulty. In the most favorable case it amounts to only about 5 percent of the total diffuse scattering level so that its evaluation by recognition of the form factor angular dependence is at most uncertain if not impossible. On the other hand, the ferromagnetic disorder scattering is sensitive to the application of an external magnetic field. Therefore, a determination of the diffuse level when the scattering sample is magnetized in different directions with respect to the scattering vector can yield the magnetic disorder scattering. None of the other diffuse scattering components mentioned above will vary with sample magnetization. In practice it has been found most convenient to compare the diffuse level with the sample unmagnetized, for which expression (1) is applicable, to that obtained when the sample is magnetized parallel to the scattering vector, in which case the magnetic scattering amplitudes go to zero and there is no magnetic disorder scattering. A correction must be applied to this change in intensity because of the altered transparency of a scattering sample when magnetized, an effect commonly known as the single transmission effect. This correction is determined for each sample and is generally small for the substances reported here.

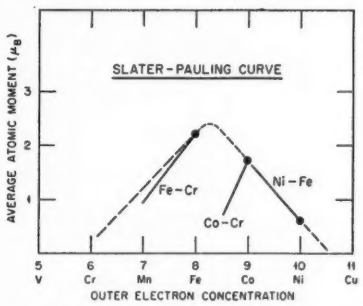


FIG. 1. Slater-Pauling curve relating the average magnetic moment to the outer electron concentration. The particular alloy systems indicated here are discussed in the text.

Ferromagnetic Superstructure Scattering

For the ordered ferromagnetic alloys, the magnetic scattering effects will be concentrated in the coherent reflections both for the normal lattice reflections and for the superstructure reflections. Again it is necessary to separate the magnetic component of the scattering from the nuclear component, and this can be done either by recognition of the magnetic form-factor portion or by determining the effect of an external magnetic field on the coherent reflections. The cross section for magnetic scattering in the superstructure reflections will be of the same form as Eq. (1) and hence the same analysis of data is possible as was performed for the disordered state.

In treating the magnetic superstructure data it is necessary to know the degree of long-range order present in the ordered sample. Usually, the long-range order parameter can be evaluated from the nuclear scattering components of the reflections and, if necessary, a suitable correction can then be applied to the magnetic intensity. This procedure is not feasible in the case of ordered Ni_3Fe since the nuclear superstructure effects are very weak. Therefore, other evidence for the degree of long-range order must be invoked.

EXPERIMENTAL RESULTS

Fe–Cr Series

Three members in this series which possessed 15.2, 29.1, and 46.4 atomic percent Cr were prepared and examined along with pure iron. Electrolytic iron bar and electrolytic chromium chips were melted inductively in a vacuum furnace and cast into bar form. Following this the sample stock was given an homogenizing heat treatment at 1300°C in a 95 percent He–5 percent H_2 atmosphere for three hours after which the sample was filed with a tungsten carbide file in a filing machine. The resulting filings sieved through 100-mesh screen were used in the study to insure randomness of crystal-

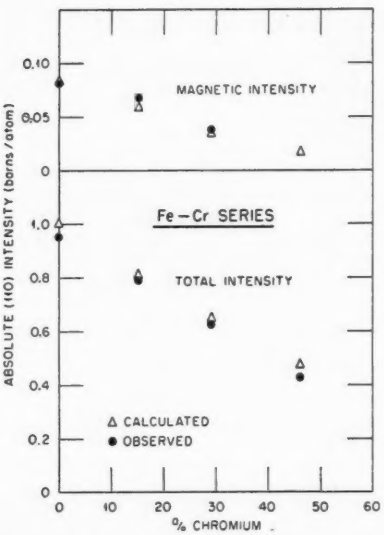


FIG. 2. Comparison between the observed and calculated values of the total intensity and the magnetic intensity in the (110) reflection for the Fe–Cr series of alloys.

lite orientation in the diffraction specimen. Chemical and spectrographic analysis were obtained on the final product of filings.

The diffraction patterns for all of these samples showed only characteristic body-centered cubic reflections with no evidence for either short-range or long-range order. Since the nuclear scattering amplitudes of Fe and Cr are rather favorable for exhibiting residual order, this can be considered good evidence of the disordered state of these preparations. In the lower portion of Fig. 2 is shown the variation of the integrated intensity of the (110) reflection, expressed in absolute units of differential scattering cross section, for the four members of the series. Along with the observed values are shown those calculated using accepted values for the nuclear scattering cross sections and the agreement can be considered satisfactory.

In the upper portion of Fig. 2 is shown the variation of the magnetic scattering cross section as seen in the (110) reflection for the different alloys. This was determined by measurement of the (110) intensity change upon magnetization of the sample in a direction parallel to the scattering vector. Also shown are the values calculated using expression (2) with average magnetic moment values obtained from magnetization studies.[4] Agreement between the observed and calculated values indicate merely that the average magnetic moment as seen by neutron scattering is consistent with that obtained in the purely magnetic type of investigation, namely the measurement of saturation magnetization.

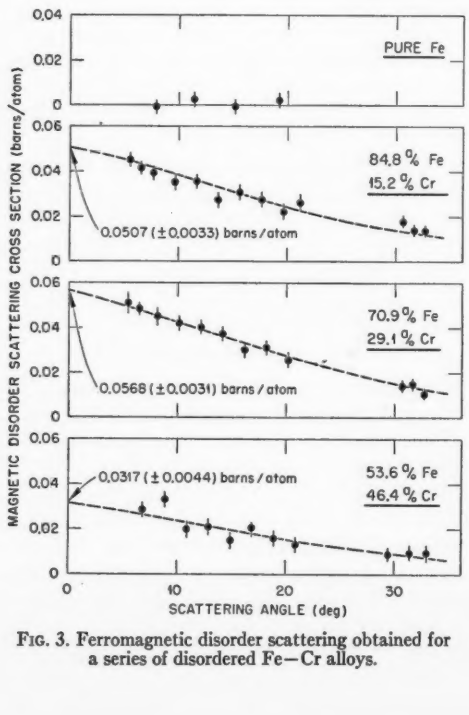


FIG. 3. Ferromagnetic disorder scattering obtained for a series of disordered Fe–Cr alloys.

TABLE I. Magnetic moment data for Fe–Cr disordered alloys as obtained from neutron scattering and magnetization experiments.

	$\mu_1-\mu_2$ (μ_B)	$\bar{\mu}$ (μ_B)
Fe	0	2.22
84.8 Fe–15.2 Cr	2.85±0.10	1.81
70.9 Fe–29.1 Cr	2.38±0.07	1.44
53.6 Fe–46.4 Cr	1.62±0.11	1.00

Of principal interest in the diffraction patterns for these alloys is the appearance of ferromagnetic disorder scattering as detected by the change in diffuse scattering intensity when the sample is magnetized parallel to the scattering vector. The observed change in intensity has been corrected for the magnetic transparency effect and placed on an absolute cross section scale with the results shown in Fig. 3. The magnetic disorder scattering for pure iron is seen to be very much smaller than that encountered in the alloys. Drawn through the various sets of data is the magnetic form factor calculated by Steinberger and Wick[5] for metallic iron and the close agreement with the observations suggest not only the applicability of this form factor but the presence of complete disorder in the alloy. The intercept of the matched form factor curves at zero scattering angle can then be used to evaluate the difference in the magnetic moments of the iron and chromium atoms and these values for the different alloys are summarized in Table

[4] The average magnetic moments for all of the alloys to be discussed in this report can be found in R. M. Bozorth, *Ferromagnetism* (D. Van Nostrand Company, Inc., New York, 1951).

[5] J. Steinberger and G. C. Wick, Phys. Rev. 76, 994 (1949).

I. Also shown in the table are values for the average magnetic moment $\bar{\mu}$ determined from magnetization experiments.

Since the algebraic sign associated with the difference values in Table I is undetermined in the present experiment, alternative sets of solutions of the individual moments are possible and these are shown graphically in Fig. 4. The size of the experimental points as shown in the figure is a measure of their accuracy. It is to be noticed that the iron moment falls as chromium is added with either of the two solutions. On the other hand, the possible chromium moment values are widely different, one being large and positive and the other small and negative in the initial concentration region. At the 50–50 concentration the solutions converge and it is impossible to distinguish chromium values from iron values.

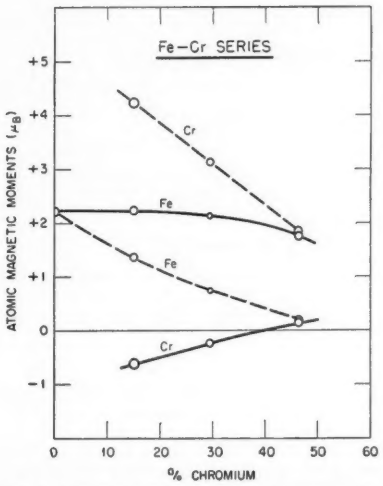


FIG. 4. Variation of the Fe and Cr atomic magnetic moments with atomic composition in Fe–Cr alloys. Alternative solutions are represented by the points connected either by the solid lines or by the dashed lines. The size of the experimental points is a measure of their accuracy.

Ni–Fe Series

Three members of this series containing 49.9, 60.1, and 74.3 atomic percent nickel were prepared and studied. These were cast from inductively heated melts of the electrolytically pure elements, homogenized at 1300°C in a 95 percent helium—5 percent hydrogen atmosphere and then filed into a 100-mesh sample preparation. In this state all of the alloys showed the face-centered cubic (f.c.c.) structure with precision lattice constant values and neutron intensities characteristic of a disordered solid solution. Additionally, a portion of the 74.3 percent nickel alloy was given a long-time heat treatment to invoke the Ni_3Fe ordered structure and this will be discussed in the following section.

TABLE II. Magnetic moment data for Ni–Fe disordered alloys as obtained from neutron scattering and magnetization experiments.

	$\mu_1-\mu_2$ (μ_B)	$\bar{\mu}$ (μ_B)
74.3 Ni–25.7 Fe	2.31±0.22	1.19
60.1 Ni–39.9 Fe	2.06±0.21	1.47
49.9 Ni–50.1 Fe	1.93±0.20	1.64

The ferromagnetic disorder scattering for these alloys is shown in Fig. 5 and again the Steinberger-Wick form factor has been matched to the various groups of data with intercept values as given in the figure. Table II summarizes the magnetic moment differences as calculated from the disorder scattering and also the average magnetic moments taken from magnetization data. The alternative sets of solutions for the individual moments have been evaluated and these are shown in graphical form in Fig. 6 with the size of the points representing their accuracy. Either set of solutions indicates that the nickel moment increases when iron is added, but the possible iron moments differ widely in the two solutions.

Ordered Ni_3Fe

It has been well established that an alloy of this composition can be ordered with a structure similar to that of Cu_3Au in which the nickel atoms are located at the face-centered positions and the iron atoms at the

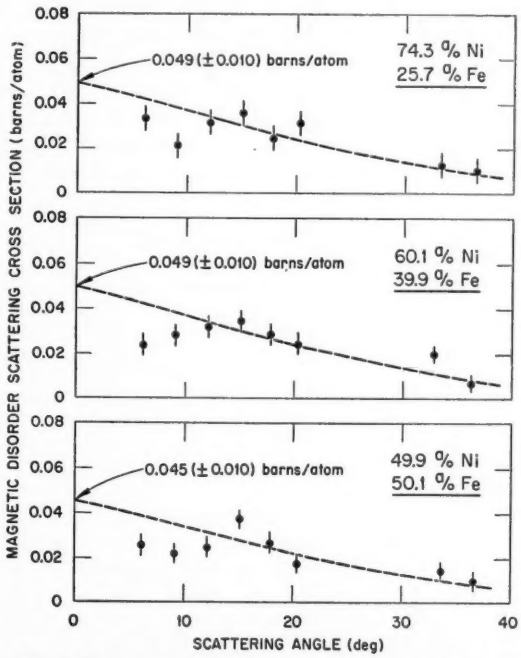


FIG. 5. Ferromagnetic disorder scattering obtained for a series of disordered Ni–Fe alloys. The decrease in scattering at small angles may indicate the presence of short-range magnetic order in the samples.

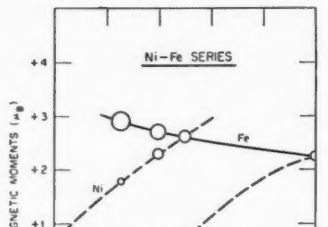


FIG. 6. Variation of the Ni and Fe atomic magnetic moments with atomic composition in disordered Ni–Fe alloys. Alternative solutions are represented by the points connected either by the solid lines or by the dashed lines. The size of the experimental points is a measure of their accuracy.

cube corners. Pronounced differences in magnetic anisotropy, permeability, and other properties have been found for the ordered and disordered preparations, and these have been used to establish an ordering temperature of 490°C. The usual diffraction technique for determining the state of long range order is not too successfully applied for this case because of the equivalence of the scattering amplitudes of Fe and Ni both for x-rays and neutrons.

Two different samples, (A) and (B), from different melts and with different ordering heat treatments have been studied. Sample (A), which contained 74.0 percent Ni, was heated in vacuum at 484°C for 200 hours after which it was cooled at the average rate of 2° every five hours to 350°C before the final cooling. Sample (B)

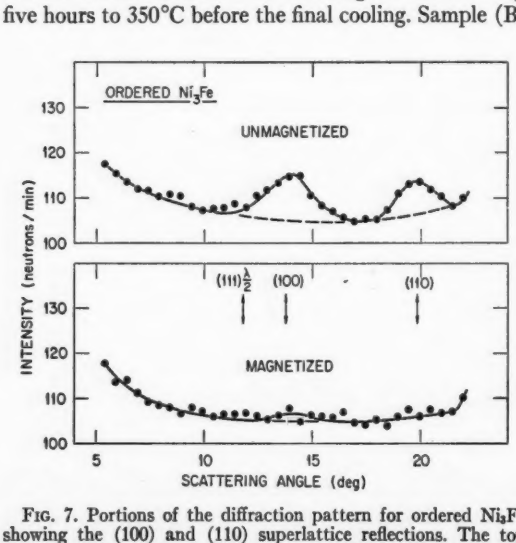


FIG. 7. Portions of the diffraction pattern for ordered Ni_3Fe showing the (100) and (110) superlattice reflections. The top pattern contains both nuclear and magnetic scattering whereas the bottom pattern contains only nuclear scattering.

which was part of the sample containing 74.3 percent Ni mentioned in the previous section, was given a much more elaborate vacuum heat treatment. The sample was first heated through the ordering temperature to 600°C for about an hour before the start of the ordering treatment. The temperature was then lowered to 500°C for 24 hours, after which it was decreased in steps of 25°, and each temperature step was held for about twice the time of the previous temperature. This procedure was followed through 375°C and required about ten weeks to complete. The superstructure lines in the diffraction pattern indicative of ordering were very weak relative to the normal f.c.c. lines. Typical portions of the patterns are shown in Fig. 7, which illustrates the (100) and (110) superstructure lines as obtained with an unmagnetized and a magnetized sample. As seen in the figure, most of the superstructure intensity is field sensitive, and this implies a *magnetic superlattice* as well as a chemical superlattice. From such studies the absolute intensity of the magnetic superstructure lines has been determined for several reflections and these results are illustrated in Fig. 8. Quite good agreement is to be noted in the data for the two samples, and the curve represents the best match of the Steinberger-Wick form factor to the data.

Before using the intercept value of 0.221 barns per Ni_3Fe molecule as the magnetic scattering cross section for the superstructure lines, it is necessary to establish the long-range order parameter. Since the nuclear scattering amplitudes are so similar for this case, it is impossible to make this determination in a direct manner by using the nuclear superstructure intensity and, therefore, indirect evidence must be invoked. It is known that the degree of order must be high, particularly for Sample (B), because of the absence of ferromagnetic disorder scattering in the diffraction pattern in contrast to that found in the disordered sample. In studies of other physical properties which are sensitive to the state of order, equivalent thermal ordering treatment has produced saturation in these properties. The agreement in the scattering data for our Samples (A) and (B) also supports a high saturation value for the degree of order. Moreover, in similar studies on Ni_3Mn (to be discussed in a later section) which orders in an equivalent manner and with closely the same ordering temperature, a long-time ordering treatment similar to that given Sample (B) resulted in a long-range order parameter of 0.905 as measured easily in the diffraction pattern. Accordingly, in the interpretation of the Ni_3Fe magnetic reflections, a long-range order parameter of 0.9 was assumed and this was considered significant to within 10 percent.

After correcting the magnetic superstructure intensity for this lack of complete order, the difference between the iron and nickel moments was evaluated as

$$\mu_{Fe}-\mu_{Ni}=\pm 2.35(\pm 0.20)\mu_B.$$

Combining this with the magnetization moment of $1.21\mu_B$ leads to the individual moments:

$$\mu_{Fe}=+2.97(\pm0.15)\mu_B, \quad \mu_{Ni}=+0.62(\pm0.05)\mu_B,$$

or

$$\mu_{Fe}=-0.55(\pm0.15)\mu_B, \quad \mu_{Ni}=+1.79(\pm0.05)\mu_B,$$

as the alternative sets of solutions. The largest part of the above values of uncertainty arises from the lack of precise information about the state of order. Very interestingly the moment difference and the individual moments for ordered Ni_3Fe agree very closely to the equivalent moments as obtained for the disordered sample of the same composition.

Co–Cr Series

Two alloys in this series were prepared containing 9.0 and 13.6 atomic percent chromium. Cobalt rondels of purity 99.3 percent and electrolytic chromium were used as starting materials and subsequent spectroscopic and chemical analysis showed about 0.5 percent nickel impurity. In the filed condition the diffraction patterns showed only the hexagonal structure lines. The ferromagnetic disorder scattering was very small for both samples amounting to only $0.0037(\pm0.0012)\times10^{-2}$ cm²/atom and $0.0047(\pm0.0016)\times10^{-24}$ cm²/atom for the 9.0 and 13.6 percent chromium alloys, respectively.

TABLE III. Magnetic moment data for Co–Cr disordered alloys as obtained from neutron scattering and magnetization experiments

	$\mu_1-\mu_2$ (μ_B)	$\bar{\mu}$ (μ_B)
91.0 Co– 9.0 Cr	0.96±0.15	1.18
86.4 Co–13.6 Cr	0.91±0.15	0.77

Table III summarizes the moment difference and average moment for the alloys and Fig. 9 illustrates the alternative sets of individual moments as calculated from the data. Of most interest here is the fact that the cobalt magnetic moment falls quite rapidly with chromium addition and that the chromium moments are all positive relative to the cobalt moments. The precipitous drop in saturation magnetization for Co–Cr as illustrated in Fig. 1 has been frequently ascribed by various authors to an antiparallel alignment of a large chromium moment in the alloy. It is seen that this is not the distinguishing feature of the magnetic structure but rather the rapid fall in the cobalt moment itself is responsible for the observed magnetic properties. If a negative chromium moment of $5\mu_B$ were present in the alloy, the disorder scattering should have been about forty times larger than was observed.

Ordered Ni_3Mn

This alloy exhibits magnetic properties which are very sensitive to the state of long range order present in the alloy. The ordered alloy is ferromagnetic with an average magnetic moment approaching $1.02\mu_B$ whereas the disordered alloy is essentially paramagnetic at room temperature with possible residual traces of weak ferromagnetism dependent upon the state of disorder. Several samples of both ordered and disordered preparations have been studied with most emphasis on a sample[6] (74.0 atomic percent nickel) which was given an ordering vacuum heat treatment identical to that given Sample (B) of Ni_3Fe which was described in a previous section. Figure 10 shows the diffraction patterns obtained for this highly ordered sample and a disordered sample of unheated filings. Pronounced superstructure intensity is to be seen in the ordered sample pattern and some residual short-range order is seen in the diffuse scattering in the disordered sample pattern. The nuclear scattering amplitudes for nickel and manganese are quite favorable for exhibiting the effects of order, and from the nuclear scattering com-

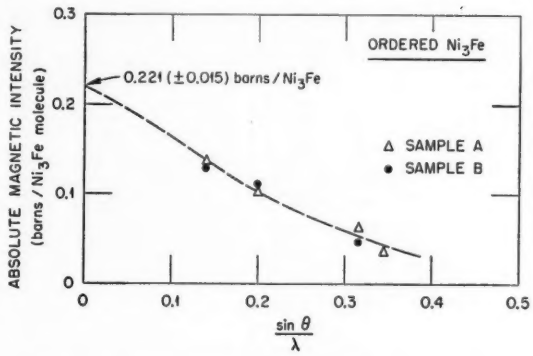


FIG. 8. Ferromagnetic scattering observed in the superlattice reflections for two samples of ordered Ni_3Fe.

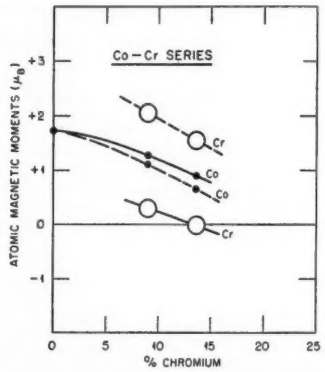


FIG. 9. Variation of the Co and Cr atomic magnetic moments with atomic composition in disordered Co–Cr alloys. Alternative solutions are represented by the points connected either by the solid lines or by the dashed lines. The size of the experimental points is a measure of their accuracy.

[6] We are indebted to Dr. Lewis R. Aronin of the Massachusetts Institute of Technology Metallurgical Project for sending us this sample of Ni_3Mn.

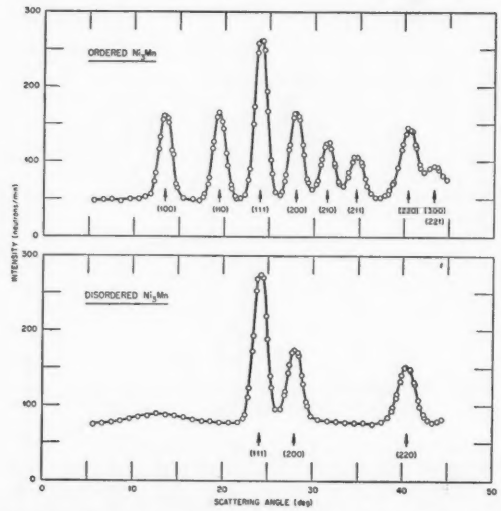


FIG. 10. Neutron diffraction patterns obtained for ordered and disordered Ni_3Mn. Residual short-range order is to be seen in the lower pattern.

ponent in the superstructure lines, the long-range order parameter is evaluated as 0.905.

The ferromagnetic components in the superstructure lines were evaluated by observing the change in intensity upon magnetization and these are shown in Fig. 11. The best magnetic form factor fit to the data yields a superstructure magnetic cross section of $0.400\ (\pm 0.061)\times 10^{-24}$ cm^2 per Ni_3Mn molecule after correction for the lack of complete order. In turn this gives $2.88\ \pm 0.20\mu_B$ as the difference between the magnetic moments of manganese and nickel in the ordered alloy. Combining this with the average ferromagnetic moment of $1.02\mu_B$ yields the alternate sets of individual moments:

$$\mu_{Mn}=+3.18(\pm 0.25)\mu_B, \quad \mu_{Ni}=+0.30(\pm 0.05)\mu_B,$$

or

$$\mu_{Mn}=-1.15(\pm 0.25)\mu_B, \quad \mu_{Ni}=+1.73(\pm 0.05)\mu_B.$$

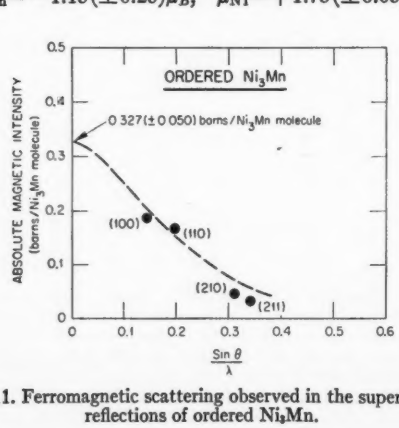


FIG. 11. Ferromagnetic scattering observed in the superlattice reflections of ordered Ni_3Mn.

Probably the first set is the correct one although the experimental data do not distinguish between the two.

A pronounced difference is to be noted in the diffuse scattering level of the patterns of Fig. 10 and this results because of the additional disorder scattering in the pattern for the disordered sample. The average nuclear disorder scattering can be calculated (this should be isotropic if no residual short-range order were present) and this accounts for most of the observed difference. Some magnetic disorder scattering modified by short-range order may very well exist in the diffuse scattering but at the present time this has not been evaluated. However, on the basis of the Ni_3Fe data, it seems unlikely that the individual moments should differ much from those in the ordered state. It can be said definitely that the disordered sample is not antiferromagnetic at room temperature.

SUMMARY

From the results presented on the above illustrative systems of alloys certain general conclusions are available. First of all, it has been shown for these alloy systems that a heterogeneous magnetic structure exists; that is, different magnetic moments are present either in a disordered or in an ordered distribution which depend upon the type of atomic distribution. Moreover, these moments are to be associated with localized atomic moments whose magnetic form factors differ little among themselves or from the theoretical calculations of the 3*d*-shell form factor as performed by Steinberger and Wick. Both the disordered and ordered magnetic scattering as seen in the diffuse scattering and superstructure lines are representative of magnetic form factor differences and the fact that this difference for Ni and Mn or Ni and Fe is significantly the same as that for Fe alone implies a closely similar *d*-shell for elements in this group.

In the case of Ni_3Fe, ordering of the lattice has been shown to have little or no effect on the component magnetic moments. This suggests that the local electron concentration of surrounding atoms has little effect on the magnetic moment and hence on the electron configuration of a central atom. Rather the over-all lattice composition seems to determine the individual moments. Changing the composition is seen to change the local moments although the significance of this is obscured because of the alternative sets of moment values consistent with the experimental data. In spite of this ambiguity it is established that the addition of Cr into either Fe or Co results in lowered moments for the solvent atoms whereas the addition of Fe into Ni elevates the Ni moment. The marked departure of the Co—Cr magnetization data from the general sequence as shown in Fig. 1 is found to result from a lowered value of the Co moment rather than an antiparallel orientation of the Cr moment.

Trapping of Minority Carriers in Silicon. I. *P*-Type Silicon

J. A. Hornbeck and J. R. Haynes
Bell Telephone Laboratories, Murray Hill, New Jersey
(Received October 11, 1954)

Experimental evidence of temporary trapping of electrons in the volume of p-type silicon in two different traps is presented. The evidence is shown to lead to a multiple trapping model the kinetics of which explains the data. The trap parameters are determined by the fit of experimental data to the theory. The trap density, energy level, cross section for capture, S, and the time which an electron spends in a trap, τ_0, are determined for both traps. For the deeper traps, $S\tau_0$, a property of the trap itself, is shown to have the probable value $0.81\pm0.25\times10^{-13}$ cm² sec with $\tau_0\cong0.3$ sec for all specimens examined. In low-resistivity crystals there is evidence for a loss mechanism for electrons (recombination) from the deeper traps at a rate found proportional to the square of the hole concentration. The deep-trap concentration is found to be roughly proportional to the sample conductivity.

INTRODUCTION

RECENT experiments[1-3] indicate that minority carriers in silicon and germanium may be trapped, that is temporarily imprisoned, at what must be special sites or imperfections in the crystal lattice. Trapping in silicon is readily observed at room temperature, whereas in germanium it is seen at ~-80°C and lower temperatures. It is the primary purpose of this paper to present (1) evidence for the existence of two sets of volume traps for electrons in p-type silicon and (2) a model for the kinetics of trapping which appears to be entirely consistent with the experimental observations. A further report, Part II, will be forthcoming shortly on hole traps in n-type silicon.

EVIDENCE FOR TRAPPING—MOBILITY EXPERIMENTS

The existence of trapping centers (other than recombination or "deathnium" centers) for minority carriers was first indicated by drift velocity experiments employing a time-of-flight technique. In this type of experiment a short pulse of minority carriers is injected by a point into a germanium or silicon rod on which is impressed a small longitudinal electric field. The carriers drift in the electric field along the rod and at the same time diffuse and also recombine with majority carriers. A second point contact placed downstream is used to measure the concentration of the minority carriers in the immediate neighborhood of the point as a function of the time and also distance from the injection point. Arrival of the minority carriers at the collector is shown schematically in Fig. 1, assuming that a short pulse of minority carriers was injected into the rod at the emitter at zero time. If the rod is germanium at room temperature, the minority carrier pulse as measured by the collector, shown by the broken line in Fig. 1, spreads out with time and distance from the emitter and decreases in amplitude with time and distance. Recombination and diffusion explain[4] these effects quantitatively. In p-type silicon at room temperature and n-type germanium *circa* -80°C and lower temperatures the pulse shape is different as shown by the solid line in Fig. 1. A straggling is observed as if some of the carriers suffer an additional time delay in their transit between the two points. The straggle can be eliminated, also as shown by the broken line in Fig. 1, by increasing sufficiently the ambient light falling on the crystal. In silicon the areas under the two curves are the same within the experimental error of measurement.

An interpretation of these results can be given qualitatively in terms of temporary traps. For low external illumination some of the carriers are caught in traps where they sit for a time and then are ejected back into the conduction stream. High external illumination, however, creates sufficient electron-hole pairs to keep the traps filled. Thus the injected minority carriers are not aware of the existence of the traps, and a pulse shape with no straggle is observed. The experiment, then, is evidence for the existence of temporary trapping centers for minority carriers in which recombination does not occur prominently, since the areas under the curves are the same. The experiment suggests that the mean time a carrier spends in a trap before release is

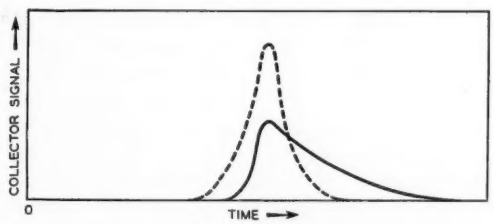


FIG. 1. Collector signal (schematic) as a function of time in the drift mobility experiment. The broken line is typical of Ge at room temperature regardless of ambient light and of Si at room temperature in strong ambient light. The solid line is typical of p-type Si in weak ambient light when some trapping of minority carriers occurs.

[1] J. R. Haynes and W. C. Westphal, Phys. Rev. **85**, 680 (1952).
[2] J. R. Haynes and J. A. Hornbeck, Phys. Rev. **90**, 152 (1953).
[3] Gebbie, Nisenoff, and Fan, Phys. Rev. **91**, 230(A) (1953).
[4] Transistor Teachers Summer School, Phys. Rev. **88**, 1368 (1952).

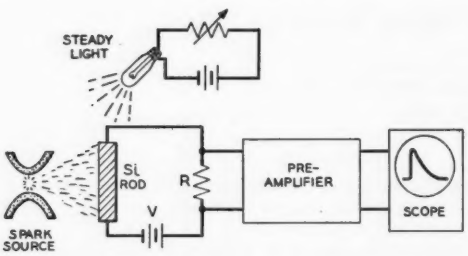


FIG. 2. Schematic diagram of circuit used to measure the lifetime of minority carriers and to study the trapping of minority carriers.

the order of tens of microseconds, i.e., the order of the minimum decay time of the straggling effect.

EVIDENCE FOR TRAPPING—LIFETIME EXPERIMENTS

Further evidence for the trapping of electrons in p-type silicon[5] has been obtained from experiments in which the lifetime of injected electrons in the material is observed directly. This experiment may be described as follows (Fig. 2).

A semiconductor sample in the form of a rod (dimensions $0.2\times0.2\times2$ cm) with plated electrodes at either end is connected in series with a resistor and a battery. A short pulse of light[6] approximately 0.2 μsec in duration illuminates the crystal. This creates very suddenly a pulse of additional electron hole pairs. The resulting change in conductivity produces a voltage change that is measured as a function of time by means of high-gain, wide-band (8-megacycle) amplifiers and an oscilloscope. Clearly, the change in voltage across the specimen, which is proportional to the increase in conductivity, will persist as long as a detectable number of added electron-hole pairs remain in the specimen. Sufficiently low electric fields are impressed on the rod so that the added pairs disappear mainly through recombination rather than drift to the electrodes. Under these circumstances the decay in photoconductivity observed is a direct measure of the apparent lifetime of the minority carriers.

If the sample is germanium at room temperature, the time constant of the decay agrees very well with the lifetime for minority carriers measured by other methods.[7] At about -80°C and lower temperatures in n-type germanium, the decay time appears to cease decreasing with temperature and, in fact, to reverse its variation with temperature. Under these conditions shining additional light on the specimen causes the decay time to shorten just as in the mobility experiment additional light causes the straggle to disappear. By using sufficient light so that the decay is independent of light intensity, the decay time has been measured as a function of temperature down to liquid nitrogen temperature. It is found to decrease monotonically with temperature and approach a constant, limiting value as predicted[8] for the variation of recombination time with temperature.

These results we interpret as additional evidence for the existence of temporary trapping centers in n-type germanium at low temperature. The implication is that the minority carriers are trapped unless sufficient electron-hole pairs are created (by the dc light) to keep the traps filled. Thus we interpret the decay time under strong external illumination as the recombination time (or lifetime) and the longer decay time as representative of some combined property of the temporary traps and recombination centers.

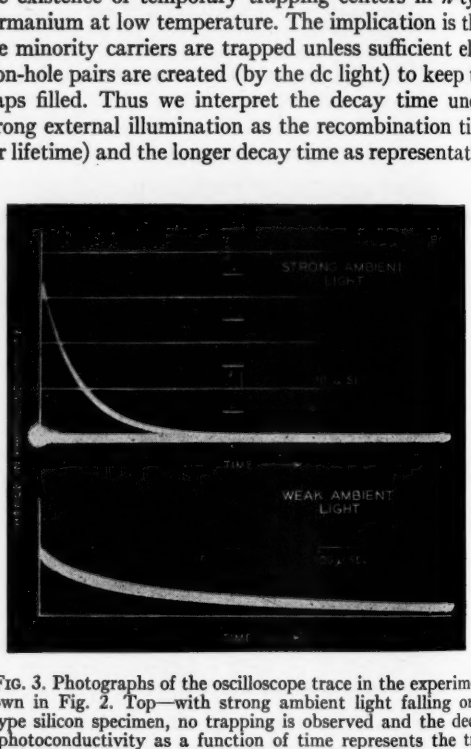


FIG. 3. Photographs of the oscilloscope trace in the experiment shown in Fig. 2. Top—with strong ambient light falling on a p-type silicon specimen, no trapping is observed and the decay in photoconductivity as a function of time represents the true lifetime of the added electrons. Bottom—without ambient light trapping occurs, and it controls the decay in photoconductivity.

At room temperature p-type silicon shows the same trapping effects in this experiment, *viz.*, a long decay time that can be changed by external dc light. Measurements of the recombination time (as described above) in this material quantitatively agree with other methods of measuring lifetime, whereas the observed decay time with weak or zero external light may be orders of magnitude longer. These effects are shown in Fig. 3. The upper part of Fig. 3 is reproduced from a photograph of the cathode ray tube trace with strong ambient illumination (light from a $1\frac{1}{2}$ volt flashlight bulb). The lower trace was photographed when light from a semi-darkened room was falling on the sample. We note that

[5] N-type silicon and germanium show similar effects.

[6] The light source is described by J. A. Hornbeck, Phys. Rev. **83**, 374 (1951).

[7] One common practice is to measure the diffusion length $l_\tau=(D\tau)^{\frac{1}{2}}$ where D is the diffusion coefficient of the minority carriers and τ is the mean lifetime before recombination. It should be emphasized that l_τ will, in general, be unaffected by temporary trapping.

[8] W. Shockley and W. T. Read, Phys. Rev. **87**, 835 (1952).

both the initial amplitude and time constant of decay can be affected by the external light.

On the basis of a trapping hypothesis it is possible to predict quantitatively the amplitude change that is observed. The experimental observation is that first the decay time increases for a while with decreasing ambient light without the initial amplitude being affected. With a further decrease in ambient light the amplitude begins to decrease and the decay time continues to increase as before. When the amplitude has decreased to one-third its initial value, a further decrease in ambient light leaves the amplitude unchanged although the decay time continues to increase.

This variation of initial deflection with dc light intensity can be explained in the following way. With all of the traps filled by the external light the initial deflection should be proportional to the change in conductivity, i.e., to the number of hole-electron pairs formed by a single light pulse and to the sum of the electron and hole mobilities: $\Delta\sigma = nq(\mu_+ + \mu_-)$, where n is number of pairs formed, q is the electronic charge, and the μ's are the mobilities of holes and electrons. As the ambient light is decreased the number of empty traps increases and the mean free time before trapping (τ_t) of an electron from the conduction band decreases. At some point τ_t will become the order of the response time of the apparatus (pulsed light source plus video amplifiers), and the initial response will fall off because those electrons that are trapped before the measuring system can "see" them give no observable conduction. For still lower light intensities (and larger number of unfilled traps), τ_t will be so short that essentially all the electrons are trapped before they are detected, and the initial amplitude should remain unchanged at that value contributed by the holes alone; since for each trapped electron there is, due to the requirement of space charge neutrality, one free hole. Thus in this limit $\Delta\sigma = nq\mu_+$. The ratio of the two limiting values of initial amplitude is $(\mu_+ + \mu_-)/\mu_+$. Rough measurements on a particular silicon specimen (223*B*) gave 3 for this ratio. If we take the values from recent mobility studies[9] $\mu_+ = 500$ cm²/volt-sec and $\mu_- = 1200$ cm²/volt-sec, we should predict a ratio of 3.4, which agrees within the accuracy of the present ratio measurement.

For this specimen the mean lifetime of electrons in the conduction band was measured as $\tau_r = 20$ μsec. Since the response time of the measuring system was $\sim 10^{-7}$ sec, we would conclude that when most of the traps are empty τ_t is less than or the order of 10^{-7} sec. Thus with the traps empty an electron in the conduction band is much more likely to be trapped in a temporary trap than to recombine, i.e., to be trapped at an ordinary recombination center. We are therefore led to the concept of multiple trapping: an electron is trapped, emitted by thermal action back into the conduction band, retrapped, re-emitted, etc., until it is finally removed from the conduction band by recombination with a free hole at a recombination (deathnium) center. The continuing increase in the decay time as the external dc illumination is decreased is further evidence for the multiple trapping process.

[9] M. B. Prince, Phys. Rev. **93**, 1204 (1954).

EXPERIMENTAL EVIDENCE FOR TWO SETS OF TRAPS

We shall present next experimental evidence for the existence of two sets of traps in p-type silicon. The evidence comes primarily from the observation of the change in photoconductivity as a function of time after external illumination is removed from a specimen. This observation is similar to the pulsed photoconductivity measurement described previously, but differs in that the light source is intense enough to saturate the traps before it is turned off.

Two variations in the experimental setup were employed to gain improved low-frequency response. In the first variation the spark light source was replaced by a battery-operated tungsten lamp and sectored disk light "chopper" as a shutter; the video amplifiers by a direct coupled preamplifier followed by either a dc oscilloscope (DuMont 304) or additional amplifiers and a Sanborn pen recorder. As connected, the high-frequency response of the setup with the oscilloscope was about 20 kc/sec and the sensitivity 10^{-3} volts/cm deflection. With a Sanborn recorder the high-frequency response was 100 cps and the detection limit about 200 μv. In the second variation for measuring very long time constants a bridge arrangement was used with an L and N dc amplifier (50-μv full scale) coupled to a G.E. pen recorder. Thus the combination of these three experimental arrangements together with the wide-band equipment described above was capable of measuring photoconductivity changes over a time range from about 10^{-7} sec to 10^3 sec or longer.

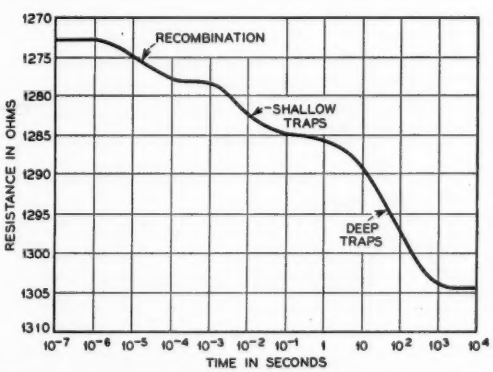


FIG. 4. Change with time of the resistance of p-type Si rod 223-*B* after a strong light has been removed from the specimen. The photoconductivity of the rod decays in three well-defined steps, the first (in sequence of time) is attributed to recombination of electrons in the conduction band; the second is attributed to the emptying of shallow traps; and the third is attributed to the emptying of deep traps.

With this apparatus the change in conductivity *versus* time shown in Fig. 4 was obtained for a particular silicon specimen (223*B*) after a strong source of light was suddenly removed from the specimen by interposing the sectored disk. Figure 5 is a reproduction of a photograph of an oscilloscope trace which shows all three decay components. Here change in conductivity is shown as a function of time. Consider the highest trace. At zero time the shutter is closed cutting off the light reaching the silicon specimen. The conductivity drops immediately because of recombination in the conduction band (this time constant is not resolved on the photograph). There follows a slower decay which takes place in tens of milliseconds. The remaining traces, photographed at the times indicated, show a much slower decay in conductivity which takes place in tens of seconds.

If the experiment is repeated successively under conditions such that the intensity of the pulsed light is decreased each time, one observes that first the amplitude of the fast component ($\tau \sim 20$ μsec) decreases and disappears; next the amplitude of the 10^{-2} sec component decreases and subsequently disappears; and then with lower initial light intensities the initial amplitude of the 260-sec component decreases.

Almost the reverse of this sequence can be brought about by starting as before but, instead of decreasing the source light intensity, gradually increasing the dc ambient light falling on the crystal. Ordinary room light is sufficient to prevent the observation of the 260-sec component. Additional light is needed to prevent detection of the 10^{-2}-sec component. There was no evidence, however, that the fast component could be eliminated.

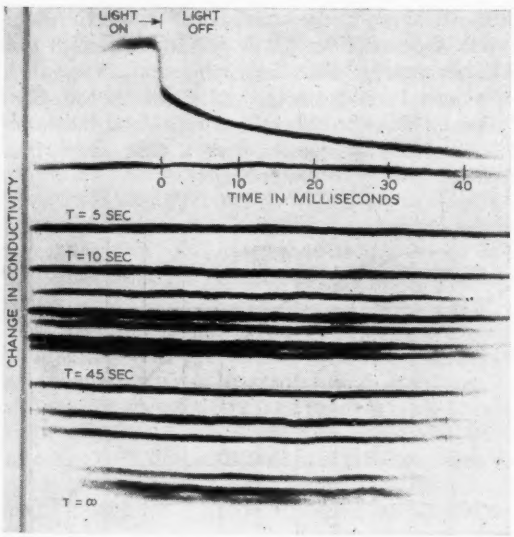


FIG. 5. Photograph of an oscilloscope pattern showing the three decay components in p-type Si (specimen 223*B*). The external light source is turned off (by a shutter) at $t=0$. The immediate decay in photoconductivity, seen in the highest trace, is attributed to recombination of electrons in the conduction band which is fast compared with the shutter speed. This is followed by a slow decay which takes place in tens of milliseconds. The remaining traces, photographed at the times indicated, show a much slower decay component which takes place in tens of seconds.

This last series of experiments demonstrates, we believe, that two distinct sets of traps for electrons are present in the silicon sample at room temperature. The fast (20-μsec) decay we have already associated with recombination from the conduction band, that is normal lifetime, before either of the sets of traps unloads appreciably. The intermediate component (10^{-2} sec) we associate with the decay of a set of relatively shallow traps. The slowest component (260 sec) we associate with the decay of a set of deeper traps.

The conclusion previously reached that multiple trapping occurs in the p-type specimen applies so far only to the shallow traps, for the effects observed in the drift mobility experiments could only be associated with the shallow traps. Under the conditions of that experiment the deep traps were filled by room light, and in any case the mean time an electron spends in a deep trap, as we shall see, is so long that it could not be observed in that experiment.

An independent experiment indicates that the deep traps also decay by a multiple trapping process. A third electrode was attached to the center of a silicon rod, and it was mounted in a bridge circuit, Fig. 6, so that the rod made up two arms of the bridge. Light was then shone for a second or two on part of the silicon rod comprising the left arm of the bridge. This unbalanced the bridge, as shown at the bottom of Fig. 6, in the direction expected for higher conductivity in that arm. In a time *the order of tens of seconds, after the light was cut off*, however, the bridge became unbalanced in the opposite direction indicating that the right arm was then of higher conductivity than the left arm. We interpret this experiment as direct evidence that electrons initially trapped in the left arm were moved due to the applied electric field into the right arm and became trapped there. Thus, multiple trapping occurs in the deep traps. This experiment also is evidence that recombination does not remove all the electrons from the deep traps.

The shape of the two trapping components of the decay curve, Fig. 4, is additional evidence of multiple trapping in both sets of traps. The components are not simple exponential functions of time as would be the case if the electrons were trapped but once.

VOLUME OR SURFACE EFFECT

By masking off the electrodes from the incident light it is rather easy to show that the delayed photoconductivity (trapping) effects are associated with the body of the semiconductor rods and not with some spurious effect at the electrodes. Also these rods show no rectification effects when the current direction is reversed.

There remains the question as to whether we are observing a surface or volume effect, or a combination of the two. It is conceivable that the entire resistivity change occurs in the narrow region within a Debye length ($\sim 10^{-4}$ cm in 20-ohm-cm silicon) of the surface. The evidence against this hypothesis is twofold. First, the decay curves (Fig. 4) are the same when the electron-hole pairs are formed (by "penetrating" light) rather uniformly throughout the volume as when they are formed preponderantly at the surface. Second, the decay curves are the same when the surface of the rod is highly polished, when it is etched, or when the specimen is sandblasted. Sandblasting increases the surface area by a factor of two or more, and the depth of the surface pits is greater than a Debye length.[10]

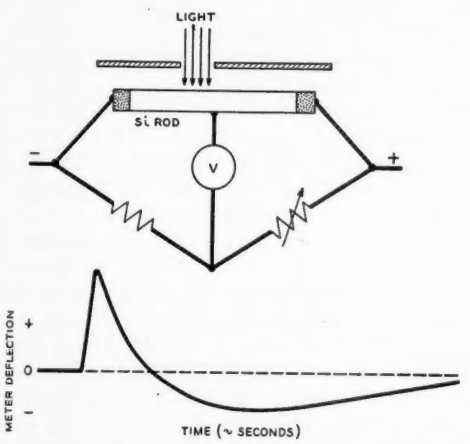


FIG. 6. Schematic representation of experiment which shows that electrons are trapped more than once in the deep traps in p-type silicon. The bridge is balanced initially with no light falling on the specimen. Light shining on part of the specimen comprising the left arm of the bridge unbalances the bridge because of electron trapping, and the meter deflects in the positive direction. The subsequent reversal of the meter which occurs tens of seconds after removal of the light shows that at a later time more electrons are trapped in the right arm than in the left arm.

The decay curves do change markedly, however, between different crystal specimens. We therefore conclude that these traps are located in the volume of the silicon.

DETERMINATION OF THE TRAPPING PARAMETERS—RESULTS

There remains to illustrate the degree of quantitative agreement between experiment and the multiple trapping model, which is treated mathematically in the Appendix. There are two sets of quantities that can be measured: amplitudes of the components and time constants. From the saturated initial amplitudes the number of normally unfilled traps is obtained, and from the decay rates come the constants of the traps.

[10] We are indebted to W. Shockley for discussions of this point.

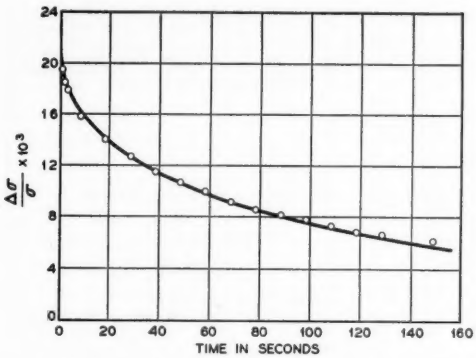


FIG. 7. The decay in photoconductivity as a function of time caused by the deep traps in p-type Si 223B. The ordinate is the fractional change in conductivity, and zero time on the abscissa is the instant the light is removed from the specimen. The open circles are experimental points, and the solid line is a fit of Eq. (10), Appendix, to these points. The values of the parameters in the equation that are determined by the fit are given in the text.

Let us first consider the deep traps in a particular p-type specimen. In Fig. 7 the open circles are experimental points taken from a record made by a pen recorder of the fractional change in conductivity of silicon specimen 223B as a function of time. This is quite obviously not a simple exponential but one in which the decay rate decreases progressively with time. According to the analysis of the multiple trapping model given in the Appendix the effective time constant of decay τ at any time t is given by Eq. (13):

$$\tau = \tau_0 + \tau_r \tau_0 NSv(1-y),$$

where τ is defined by $\tau^{-1} = -(1/y)dy/dt$, τ_0 is the mean time that an electron spends in a trap, τ_r is the mean lifetime of electrons in the conduction band, N is the density of normally unfilled traps, S is the cross section for capture of electrons in traps, v is the arithmetic mean thermal velocity, and y is the fraction of traps filled. The solid line of Fig. 7 is a fit of Eq. (13) to these experimental data. The values of the parameters determined by the fit are: density of normally unfilled deep traps, $N = 1.0 \times 10^{13}/\text{cm}^3$, the time constant at infinite time, as given by Eq. (6) of the Appendix, $\tau_\infty = \tau_r \tau_0 NSv = 260$ sec, $\tau_0 \simeq 1$ sec, and $\tau_r = 20 \times 10^{-6}$ sec. The thermal velocity v for a particle having the mass of a free electron at room temperature is 1.07×10^7 cm/sec. Substituting these values into Eq. (6) we may estimate the capture cross section of the deep traps $S \simeq 10^{-13}$ cm^2.

The data for the shallow traps in specimen 223B shown in Fig. 5 are not accurate enough to warrant a careful fit to Eq. (13). Instead we have measured directly the time constant for the shallow trap τ_∞ and the initial conductivity (when all the traps are filled). From these we find the concentration of shallow traps not normally filled $N = 2 \times 10^{12}/\text{cm}^3$ and $\tau_\infty = 10^{-2}$ sec. From the drift velocity experiment mentioned earlier

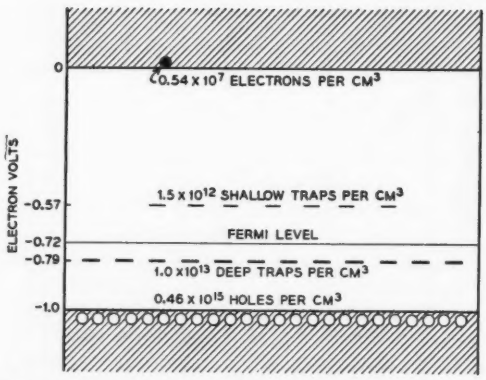


FIG. 8. Diagram depicting the number and relative energy of the traps for electrons in p-type Si 223B. Zero of energy is the bottom of the conduction band. The numbers given were computed assuming the specimen to be at room temperature.

we obtain τ_g for the shallow traps[11] $\simeq 50\times10^{-6}$ sec. Since $\tau_r=20\times10^{-6}$ sec, we find using Eq. (6) that the time before trapping in a shallow trap $\tau_t=1.2\times10^{-7}$ sec and the capture cross section of a shallow trap $S=4\times10^{-13}$ cm^2.

It is evident that the cross sections associated with the trapping of electrons in deep and shallow traps are large on an atomic scale. Wannier of these Laboratories has estimated[12] a limiting value for τ_t (not cross sections directly) by assuming that the trapping process involves emission of a phonon by a very slow electron. He finds that the free time before trapping of electrons in silicon $>10^{-9}$ sec. This limit is two orders of magnitude less than the trapping time $(NSv)^{-1}$ for either the deep or shallow traps, and therefore lends credibility to the large values obtained for their capture cross sections.

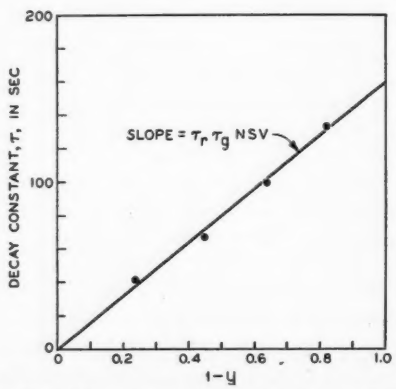


FIG. 9. Plot of the apparent time constant τ as a function of $(1-y)$, the fraction of deep traps empty (specimen VI-550-1, $\rho=17$ ohm cm). According to the simple multiple-trapping model Eq. (13), this plot should be a straight line with intercept at $y\equiv n_2/N=1$ equal to τ_g. The quantity τ_g is too small to be resolved on this time scale.

DEPTH OF TRAPS

A well-known argument involving detailed balance, based upon the assumption of thermal equilibrium, permits us to estimate the depth[13] of the traps below the conduction band. This formula[14] in a form useful for the present calculations is

$$\exp[-(E_c-E_T)/kT]=N_c\tau_g/N\tau_t.$$

Here E_c-E_T equals the energy difference between the traps and the conduction band, k is Boltzmann's constant, T is the absolute temperature, $N_c=2.41\times10^{19}/\text{cm}^3$ at room temperature is the so-called equivalent density of states in the conduction band, assuming a spherical, nondegenerate band, and τ_g, τ_t, and N refer to the traps as previously defined. Here we have ignored the statistical weight of the traps since it will not affect things much. Since closely $\tau_g/\tau_t=\tau/\tau_r$, the ratio τ_g/τ_t is known much more accurately from these experiments than either of the quantities is known by itself. Upon substituting the known quantities in the detailed balance equation we find, $E_c-E_{T_1}=0.57$ ev for the shallow traps and $E_c-E_{T_2}=0.79$ ev for the deep traps. From the measured resistivity of specimen 223B, *viz.*, $\rho=27$ ohm cm, (E_c-E_F), where E_F is the Fermi level, is ~0.72 ev. Thus the deep traps lie about $3kT$ below the Fermi level. The thermal equilibrium situation in this specimen for room temperature is given in Fig. 8. The numbers given are computed assuming that the electron mass is the mass of a free electron.

By definition, N, the number of normally empty trapping sites, is given by $N=N_0(1-f)$ where N_0 is the total number of trapping centers and

$$f=[1+\exp(E_T-E_F)/kT]^{-1}$$

is the Fermi factor. Thus the number of deep trapping centers is $\sim1.7\times10^{14}/\text{cm}^3$, whereas the number of shallow trapping centers very closely equals the number of normally empty shallow traps.

RESULTS FOR A NUMBER OF SPECIMENS AND RECOMBINATION IN THE DEEP TRAP

Since, as we have seen, the multiple-trapping model appears to agree quantitatively with experiments on specimen 223B, the question arises as to whether the analysis is generally applicable to p-type silicon. It will be shown in this section that this is the case for the

[11] The mean time in a shallow trap, τ_g, should be the decay time of the "straggle effect" in the drift velocity experiment when the electrons have a small probability of being trapped once in traveling from emitter to collector and therefore a very small or negligible change of being trapped twice. This can be arranged experimentally by having enough ambient (dc) light falling on the specimen to fill all the shallow traps, and then reduce the ambient light intensity just a little so that some trapping occurs.

[12] G. H. Wannier, Phys. Rev. 91, 207(A) (1953).

[13] Efforts to observe the trap depth by the change in photoconductivity associated with the absorption of infrared radiation have been unsuccessful. The sensitivity of the apparatus was such that the change in conductivity with light which was observed could be accounted for by a change in temperature of the silicon rod of about 0.001°C.

[14] A further discussion of the relationship and justification for its use will be found in Part II of this report.

deep traps provided that an additional loss mechanism for electrons is added to the model,[15] *viz.*, recombination of a trapped electron directly from the deep trap.

A convenient way to analyze a multiple-trapping decay curve, such as that in Fig. 7, is to plot the apparent time constant τ as a function of $(1-y)$, the fraction of traps empty. According to Eq. (13), this plot should be a straight line with intercept τ_g at $(1-y)=0$ and of slope $\tau_r\tau_g NSv$. Such a plot is given in Fig. 9 for specimen (VI-550-1) which is p-type with $\rho=17$ ohm cm. In this case the intercept τ_g is not resolved, since there is so much multiple trapping.

In higher-conductivity material straight-line plots of τ *vs* $(1-y)$ are not obtained, indicating a deviation from the model as presented. An example of this is shown in Fig. 10 for silicon specimen (VI-277), $\rho=1.9$ ohm cm. This deviation can be accounted for quantitatively by assuming that an electron in a trap has a certain probability of recombining with a hole without first returning to the conduction band. If the decay

TABLE I. Summary of data on deep traps in p-type silicon analyzed by assuming the multiple trapping model with recombination.

Sample	Conductivity (ohm cm)$^{-1}$	Trap concentration[a] cm^{-3}	τ_∞ sec	τ_r μsec	τ_b sec	$S\tau_g=\tau_\infty/\tau_r Nv$ cm^2 sec
259-*H*	1.02	1.48×10^{14}	41.6	~0.5	32	0.5×10^{-13}
323-*D*	0.625	2.83×10^{14}	516	4	78	0.43×10^{-13}
277	0.524	0.92×10^{14}	526	11	111	0.49×10^{-13}
546-1	0.14	1.88×10^{13}	435	20	∞	1.1×10^{-13}
550-1	0.0599	1.00×10^{13}	161	19	∞	0.79×10^{-13}
223-*B*	0.0373	1.00×10^{13}	260	33	∞	0.74×10^{-13}
568*A*-3	0.0369	3.82×10^{12}	150	26	∞	1.41×10^{-13}
509-2	0.021	1.35×10^{12}	0.7	1.1	∞	0.46×10^{-13}
580*A*-3	0.0178	7.1×10^{11}	23	23	∞	1.32×10^{-13}
580*A*-1	0.015	6.65×10^{11}	19	29	∞	0.91×10^{-13}

[a] Corrected for variation in mobility with conductivity.

rate from the trap through this process is $(1/\tau_b)$, then it is easily shown that the observed $\tau(y)$ is given by [see Eq. (31) of the Appendix],

$$(1/\tau)=(1/\tau_b)+(1/\tau)_{old},$$

where $(1/\tau)_{old}$ is that defined by Eq. (13). If there is sufficient multiple trapping so that τ_g is negligible compared with the other term in $(1/\tau)_{old}$, then according to the above equation a plot of $(1/\tau)$ *versus* $(1-y)^{-1}$ should yield a straight line of slope $(\tau_r\tau_g NSv)^{-1}$ and intercept at $(1-y)^{-1}=1$ of $[\tau_b^{-1}+(\tau_r\tau_g NSv)^{-1}]$. The data of Fig. 10 are replotted in this way in Fig. 11. The indication that this analysis is correct comes not only from the fact that the data plot in a straight line. In addition, the value of $S\tau_g$, which is a property of the deep trap itself, turns out to be within experimental error just that which is obtained in higher-resistivity specimens for which $1/\tau_b$ is negligible.

[15] J. A. Hornbeck and J. R. Haynes, Phys. Rev. **94**, 1437(A) (1954).

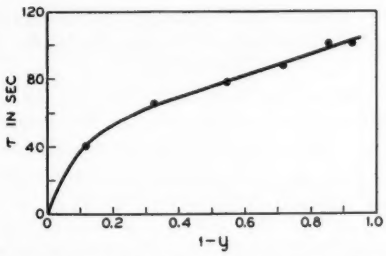


FIG. 10. Plot similar to that in Fig. 9 showing a deviation from the simple multiple-trapping model (Specimen VI-277, $\rho=1.9$ ohm/cm).

Data for a number of different specimens derived from analyses as described above are summarized[16] in Table I.

From Table I, we have computed the most probable value of $S\tau_g$ and find it equal to $(0.81\pm0.25)\times10^{-13}$ cm^2 sec, which corresponds to a single deep-trapping level 0.78 ev below the conduction band. Specimen (509-2) was carefully chosen because in it τ_∞ is the same order of magnitude as τ_g, so there is a measurable intercept when its decay curve is plotted as in Fig. 9. From this we obtain $\tau_g\simeq0.3$ sec, whence $S\simeq3\times10^{-13}$ cm^2.

The consistency of the results in Table I gives us some confidence in the analysis, and we therefore feel justified in pursuing further recombination in the deep trap as characterized by τ_b. These data from Table I together with data for higher-conductivity specimens with single exponential decays (because τ_b predominates) are shown in Fig. 12, where the rate $1/\tau_b$ is plotted as a function of the hole concentration. Within the experimental error, $1/\tau_b$ is found to be proportional to the square of the majority carrier concentration, and the rate of the recombination is low as if "forbidden" by a selection rule. We do not know what the recombination mechanism is, and we shall not speculate on it at this time.

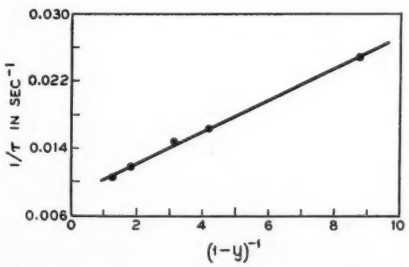


FIG. 11. Replot of the data, Fig. 10, as suggested by Eq. (31) which is derived on the assumption that, in addition to simple multiple-trapping, recombination in the trap also occurs.

[16] In studying the deep traps, it is particularly important to keep stray light from falling on the sample. Not only does this affect the measurement of the trap density, but also the values of the trapping parameters. See Appendix, Case IV.

CORRELATION BETWEEN DEEP-TRAP CONCENTRATION AND SAMPLE CONDUCTIVITY

A correlation has been found between the deep-trap concentration and conductivity of the p-type silicon specimens. This is shown in Fig. 13. A semiquantitative interpretation of these results is that the material has nearly a constant number ($\sim 10^{15}$ cm^{-3}) of trapping centers and that the number of "cocked" or empty traps varies with the position of the Fermi level, which itself is a function of conductivity. Recently[17] it has been found that the trap concentration can be reduced if when the single crystal is grown, the seed is not rotated. We have no explanation for this effect, nor do we know at this time what causes the deep or shallow traps.

ACKNOWLEDGMENTS

We wish to thank many of our colleagues for considerable help during the course of this investigation. These include H. R. Moore, who guided the instrumentation, E. Buehler, who supplied silicon crystals, W. C. Westphal, who assisted us generally, W. Shockley, C. Herring, G. H. Wannier, and P. A. Wolff, who supplied valuable discussions, and M. D. Underwood, who made some of the measurements of trap concentrations.

APPENDIX: TRAPPING KINETICS

With the multiple-trapping model suggested by experiment, we shall now consider the parameters that enter into this volume process and deduce for comparison with experiment the mathematical form of the decay curves on the basis of the model.[18] In these calculations we shall be concerned with deviations from thermal equilibrium.

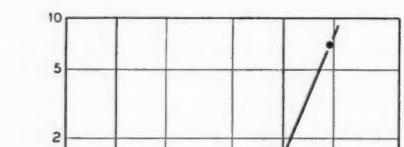


FIG. 12. Plot of the rate of recombination in the deep traps, $1/\tau_b$, as a function of the majority carrier concentration. Within the experimental error, the slope of the line equals 2.

[17] Hannay, Haynes, and Shulman, Phys. Rev. **96**, 833 (1954).

Let n = the excess density of electrons in the conduction band, n_1 = the excess density of electrons in traps, $r \equiv 1/\tau_r$ = rate of recombination of excess electrons in the conduction band, $g \equiv 1/\tau_g$ = rate of generation of electrons from traps, S = cross section for capture of electrons in traps, N = density of normally unfilled traps, v = arithmetic mean thermal velocity, l = rate of generation of electron-hole pairs per cm^3 by external light, and $NSv \equiv 1/\tau_t$ = rate of trapping of electrons when all traps are empty. With these definitions we may immediately write down the differential equations which follow for the case in which there is a single set of volume traps. We shall see later that it is unnecessary to consider in detail the case of two sets of traps because the time constants of the traps in p-type silicon usually are so well separated.

$$dn/dt = l - rn + gn_1 - n(N - n_1)Sv,$$
$$dn_1/dt = -gn_1 + n(N - n_1)Sv. \tag{1}$$

In all the following cases the observed change in conductivity for a p-type specimen is given by

$$\Delta\sigma = nq(\mu_- + \mu_+) + n_1 q\mu_+. \tag{2}$$

As we shall see, the first term of Eq. (2) is negligible compared to the trapping term except (a) at very high values of l and (b) for a short time (the order of the recombination time, τ_r) after the light is turned off. Thus $\Delta\sigma = n_1 q\mu_+$ and is a direct measure of n_1, the number of trapped electrons.

Case I. Steady State

In the steady state, $dn/dt = dn_1/dt = 0$, and the solution of (1) is[19]

$$n = l/r,$$
$$n_1/N = (1 + g/nSv)^{-1}. \tag{3}$$

Case II. Transient Solution for No External Light

Suppose that external light is shone on the crystal, and at $t = 0$ the light is turned off, i.e., for $t < 0$, $l = l_0$ and for $t \geq 0$, $l = 0$. We shall consider first the asymptotic form of the solution. For sufficiently large t, n_1 will become small compared to N and Eqs. (1) become

[18] H. Y. Fan, Phys. Rev. **92**, 1424 (1953), independently has considered parts of this same problem.

[19] The equivalent steady-state solution of the problem when two sets of traps are considered is $n = l/r$, $n_1/N_1 = (1 + g_1/nS_1v)^{-1}$, and $n_2/N_2 = (1 + g_2/nS_2v)^{-1}$. Here the subscripts (1) refer to one set of traps and the subscripts (2) to the same quantities for the second set of traps. Physically, the important point here is that the fraction of the traps of either set filled in the steady state by light is independent of the existence of the other set.

linear, i.e., for large t

$$dn/dt \simeq -rn + gn_1 - (NSv)n,$$
$$dn_1/dt \simeq -gn_1 + (NSv)n. \quad (4)$$

It is easy to show that the solution of (4) is given to a good approximation by[20]

$$n \propto \exp(-t/\tau_\infty),$$
$$n_1 \propto \exp(-t/\tau_\infty), \quad (5)$$
$$\tau_\infty = \tau_r + \tau_g + \tau_r\tau_g NSv.$$

Thus the decay curve approaches asymptotically a single exponential of time constant τ_∞. It turns out that in almost all cases for p-type silicon $(\tau_r+\tau_g) \ll \tau_r\tau_g NSv$, so that[21]

$$\tau_\infty \simeq \tau_r\tau_g NSv \equiv \tau_r\tau_g/\tau_t, \quad (6)$$

where $\tau_t = (NSv)^{-1}$ is the mean free time an electron spends in the conduction band before trapping when all the traps are empty.

Because the recombination rate is small compared to NSv, as approximate solution of (1) for $l=0$ can be obtained. This solution can easily be more exact than the experimental error in measurement. We shall normalize Eqs. (1) in order to illustrate the extent of the approximation. We define some dimensionless parameters as follows:

$$T = (rg/NSv)t, \quad G = g/NSv, \quad R = r/NSv,$$
$$x = nSv/g, \quad y = n_1/N. \quad (7)$$

Then for $l=0$ Eqs. (1) become

$$Gdx/dt = -x + R^{-1}[y - x(1-y)], \quad (8a)$$
$$dy/dT = -R^{-1}[y - x(1-y)]. \quad (8b)$$

As we shall demonstrate later, G is the order of 10^{-6} for the deep traps and 2×10^{-3} for the shallow traps, whereas R^{-1} may be greater than ten and usually $|Gdx/dT| \ll |dy/dT|$. Thus we set $Gdx/dT=0$ in Eq. (8),[22] i.e.,

$$x = y(1+R-y)^{-1}, \quad (9)$$

[20] Figure 4 helps in arriving at the equivalent solution to Eqs. (5) for a two-trap model. According to Fig. 4, the shallow traps all unload before the number of electrons in the deep traps changes appreciably. Thus we have to consider only the effect of the vacant shallow traps on the decay while the deep traps are unloading. To begin with, we don't expect much of an effect because decay times associated with the shallow traps are several orders of magnitude smaller than those associated with the deep traps. It is easy, however, to obtain an approximate solution for the two-trap case if one assumes (a) that the shallow traps are substantially empty, and (b) that the fraction of the deep traps filled is small, *viz.*, $(n_2/N_2 \ll 1)$. For this case the exponential decay time of the system $\tau_{\infty 2}$ is given by

$$\tau_{\infty 2} = \tau_r + \tau_{g1} + \tau_{g2} + \tau_r\tau_{g1}/\tau_{t1} + \tau_r\tau_{g2}/\tau_{t2}.$$

For the case in hand, the last term of this expression turns out usually to be at least 100 times as large as any of the other terms, and it comes from a single-trap model.

[21] This formula is given by J. R. Haynes and J. A. Hornbeck, Phys. Rev. **90**, 152 (1953).

[22] We are indebted to Dr. R. W. Hamming for pointing out this approximation. A more complete justification for this approximation is given in this Appendix, Case III.

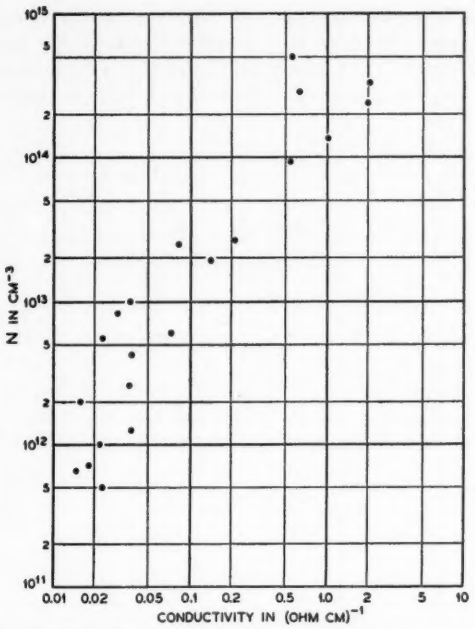


FIG. 13. Plot of the number of empty (deep) traps *vs* sample conductivity showing a rough correlation which can be accounted for largely by the position of the Fermi level if the number of trapping centers in each specimen is roughly constant.

and integrate. This yields, for the initial condition that $y=y_0$ at $t=0$,

$$\exp-[T/(1+R)] = (y/y_0)\exp[(y_0-y)/(1+R)]. \quad (10)$$

According to Eq. (10), for sufficiently large t, y decays exponentially with a time constant τ_∞ given by

$$\tau_\infty = \tau_g + \tau_r\tau_g NSv, \quad (11)$$

which differs only by the negligible quantity τ_r from Eq. (5).

We wish to emphasize particularly one property of the multiple-trapping equation (10) above which is quite useful in reducing experimental data. We first define the apparent time constant of decay, τ, at any time t to be

$$1/\tau = -(1/y)dy/dt. \quad (12)$$

Now using Eq. (10), we find

$$\tau = \tau_g + \tau_r\tau_g NSv(1-y). \quad (13)$$

In practice, from an experimental decay curve $y(t)$ τ is obtained and plotted *vs* $(1-y)$, the fraction of traps empty. This plot (in the absence of any recombination directly in the trap—see Appendix Case V) is a straight line of slope $\tau_r\tau_g NSv$ and intercept at $t=0$ and $y=1$, of τ_g.

Case III. Solution Near Zero Time and Value of $x = nSv/g$

It has been stated without proof that the contribution to the added conductivity of the conduction-band elec-

trons is negligible compared with the contribution of the holes associated with trapped electrons. We shall now examine the range of validity of this statement and the approximation made in the last section leading to Eq. (10).

If external light has been shining on the crystal long enough for a steady state to be established, then from Eqs. (3),

$$(n_0/N)\equiv Gx_0=l/Nr,\qquad (n_{10}/N)\equiv y_0=(1+1/x_0)^{-1}. \tag{14}$$

If the external light is turned off at $t=0$, then for all time both dx/dT and dy/dT are negative, i.e., both the concentration of electrons in the conduction band and the concentration of trapped electrons decay with time. This is shown in the second part of this section. From Eqs. (8), a direct consequence of this is that for all time,

$$-x+R^{-1}[y-x(1-y)]<0,\qquad -R^{-1}[y-x(1-y)]<0. \tag{15}$$

The relationships (15) may be expressed as the following inequality:

$$\frac{y}{1-y}>x>\frac{y}{1-y+R}. \tag{16}$$

As will be seen, for both deep and shallow traps $R=r/NSv\sim 0.005$ in the silicon specimen 223*B*. Thus for time greater than that at which $(1-y)$ is several times R, i.e., $(1-y)\simeq 0.02$, there is no appreciable difference between $(1-y)$ and $(1-y+R)$, and quite accurately $x=y(1-y+R)^{-1}=y(1-y)^{-1}$. In this region Eqs. (9) and (10) are the solution. The expression for x can be written in the following form, making use of the inequality (16)

$$n/N\leq Gy/1-y.$$

For the shallow traps $G\simeq 2\times 10^{-8}$ and for the deep traps $G\simeq 10^{-7}$. Thus when $(1-y)\geq 0.02$, $n/n_1<0.1$ for the shallow traps and for the deep traps, $n/n_1<0.5\times 10^{-4}$; when $(1-y)\geq 0.1$, $n/n_1<0.02$ for the shallow traps. We conclude, then, that over almost the entire decay curve for either deep or shallow traps $n_1\gg n$ and $\Delta\sigma\simeq n_1 q\mu_+$.

Next we shall consider the solution to Eqs. (8) for the shallow traps in the region near $t=0$, when the approximate solution given by Eqs. (9) and (10) is invalid. Equation (8) above is generally satisfactory for the deep traps. For time in the immediate vicinity of $t=0$ we expect that $dx/dT\gg dy/dT$. If $dy/dT\simeq 0$, then from Eq. (8a) $Gdx/dT\simeq -x$. This suggests that we try a solution of the form

$$x=x_0\exp(-T/G)+f(T), \tag{17}$$

whence

$$Gdx/dT=x_0\exp(-T/G)+df/dT$$

and

$$-Rdy/dT=y-(1-y)[x_0\exp(-T/G)+f]. \tag{18}$$

At $T=0$, $f(T)=0$; in some region $x_0e^{-T/G}\gg f$. Let us solve Eq. (18) in this region by neglecting f, i.e., letting

$$T/G\equiv T_r\equiv rt,\text{ and } G/R\equiv\lambda,\qquad -\lambda^{-1}dy/dT_r\simeq y-(1-y)x_0\exp(-T_r). \tag{19}$$

It follows directly that

$$y\simeq\lambda x_0F^{-1}(T_r)\int e^{-T_r}F(T_r)dT_r+AF^{-1}(T_r), \tag{20}$$

where A is a constant of integration and $F(T_r)$ is

$$F(T_r)\equiv\exp[\lambda(T_r-x_0e^{-T_r})]. \tag{21}$$

Successive integration by parts expands Eq. (20) into the following asymptotic series

$$y\simeq 1-(e^{T_r}/x_0)+(1+\lambda^{-1})(e^{T_r}/x_0)^2-(1+\lambda^{-1})(1+2\lambda^{-1})(e^{T_r}/x_0)^3+\cdots+(-1)^n\times\left(1+\frac{n-1}{\lambda}\right)\left(1+\frac{n-2}{\lambda}\right)\left(1+\frac{n-3}{\lambda}\right)\cdots\times(1)(e^{T_r}/x_0)^n+\cdots+AF^{-1}. \tag{22}$$

The solution comprising the first two terms of Eq. (22), it turns out, is valid for large enough values of T_r so that it overlaps the region in which the solution given by Eqs. (9) and (10) is valid. We must, however, adjust the arbitrary constant in Eq. (10) so that the two solutions fit together continuously. Our final result is for $0<T_r\leq T_{rc}$, where $T_r=T/G$,

$$x=x_0e^{-T_r},\qquad y=1-x_0^{-1}e^{T_r}. \tag{23}$$

For $GT_{rc}=T_c\leq T<\infty$,

$$x=y(1-y+R)^{-1},\qquad \frac{y}{y_c}\exp\left(\frac{y_c-y}{1+R}\right)=\exp\left(\frac{-(T-T_c)}{1+R}\right), \tag{24}$$

where

$$T_{rc}=\ln(x_0\sqrt{G}),\qquad x_c=1/\sqrt{G},\qquad y_c=1-\sqrt{G}. \tag{25}$$

The critical values, *viz.*, x_c, y_c, and T_c, are chosen such that to within a negligible error both the slopes and absolute magnitudes of the two expressions (23) and (24) for y are equal.

We now can see the range in which the series (22) needs to be valid: The relations (25) state that the maximum value of $(e^{T_r}/x_0)=\sqrt{G}$, and this is independent of x_0 and, therefore, the initial light intensity except that x_0 must be greater than x_c. If $x_0<x_c$, the solution Eq. (10) holds for all T; this is the case for the deep traps and any reasonable intensity of light prior to $T=0$. Although the series (22) eventually diverges if too many terms are retained, the error incurred by

keeping only the first two terms is less than the absolute magnitude of the next term.[23] By a similar argument for the case $T_r=0$, it can be shown that the term AF^{-1} is negligible, i.e., $A=0$. Since F^{-1} decreases with increasing T_r, AF^{-1} is negligible at all times in which we are interested.

We should point out that at $T_r=0$, Eqs. (23) give

$$\left.\frac{dy/dT_r}{dx/dT_r}\right|_0=\frac{1}{x_0^2},$$

which for large x_0 is consistent with our previous assumption in setting up this solution. Note also that at $T_r=0$, both derivatives are negative, which confirms an earlier assertion. It is easy to show, also, that $f(T)$ defined by Eq. (17) is very small and negligible.

Case IV. Transient Solution with Steady External Light

Another feature of the multiple-trapping model that can be tested by experiment is the variation of the decay curves with arbitrary ambient illumination. This in particular emphasizes the nonlinear character of the decay and illustrates the importance of stray light in affecting the values of the trap parameters as deduced from experiment.

Equation (14) states that steady illumination l_0 maintains an additional number of electrons n_0 in the conduction band and n_{10} in traps, where

$$n_0=l_0/r,$$
$$n_{10}/N=[1+g/n_0Sv]^{-1}. \tag{14}$$

Let us now consider deviations from this steady state due to a second light source which is turned off at $t=0$. The deviations u and u_1 in n and n_1, respectively are defined by

$$u=n-n_0,$$
$$u_1=n_1-n_{10}. \tag{26}$$

On substitution of (26) into (1) we obtain

$$du/dt=-ru+(g+n_0Sv)u_1-u[(N-n_{10})-u_1]Sv,$$
$$du_1/dt=-(g+n_0Sv)n_1+u[(N-n_{10})-n_1]Sv. \tag{27}$$

Clearly Eqs. (27) have the same form as the normalized Eqs. (8) if we replace the parameters as indicated below:

$$N\rightarrow N(1-n_{10}/N),$$
$$g\rightarrow g(1-n_{10}/N)^{-1},$$
$$x\rightarrow\frac{uSv}{g}(1-n_{10}/N), \tag{28}$$
$$y\rightarrow u_1[N(1-n_{10}/N)]^{-1},$$
$$T\rightarrow T(1-n_{10}/N)^{-2}.$$

[23] We are indebted to Dr. R. W. Hamming for a discussion of asymptotic series.

The relations (28) predict, for example, that the asymptotic decay constant τ_∞ will vary as the square of the fraction of traps that are not filled by the steady illumination. In the case of the deep traps, this has been checked quantitatively.

The kinetics associated with filling the traps, i.e., the case in which the traps are empty and at $t=0$ the external light is turned on, will not be discussed here because this case is not essential in establishing the trapping model. It may be pointed out that if the probability of an electron being trapped is substantially larger than the probability that it recombine, the electrons formed by the light will tend to fill up the traps before building up the steady-state concentration in the conduction band. If, however, the recombination rate greatly exceeds the trapping rate, then after the light is turned on, the steady-state concentration in the conduction band will build up (in the order of the recombination time) before appreciable trapping occurs.[24]

Case V. Trapping Kinetics with Recombination in the Trap

We now introduce a second loss mechanism for electrons, *viz.*, recombination in (or from) the trap. For the case $l=0$ (no external light), Eqs. (1) become

$$dn/dt=-rn+gn_1-n(N-n_1)Sv,$$
$$dn_1/dt=-(g+b)n_1+n(N-n_1)Sv. \tag{29}$$

Here $b\equiv 1/\tau_b$ is the rate of recombination from the trap just as r is the rate of recombination from the conduction band. Now we make the same approximation that was used to solve Eqs. (8), and again we can handle the problem. We find

$$e^{-bt}=y^{(1+R)/(1+R+PR)}\left[\frac{1+R+PR-y}{R(1+P)}\right]^{PR/(1+R+PR)}, \tag{30}$$

for the approximate initial condition that for $t=0$, $y=1$. Here the notation is that defined by (7) with the addition that $P\equiv g/b$.

In the limit that $b\rightarrow 0$, i.e., no recombination in the trap, Eq. (30) becomes Eq. (10), as it should. In the other limit, $b\gg g$, Eq. (30) approaches the limit $y=\exp(-bt)$, as it should. This case also is realized experimentally for material of sufficiently high conductivity (see, for example, Fig. 12).

As suggested in the text, analysis for the recombination effect τ_b is facilitated by plotting from the experimental data $\tau^{-1}\equiv-(1/y)(dy/dt)$ *vs* $(1-y)^{-1}$. From Eq. (30) we find

$$1/\tau\equiv-(1/y)(dy/dt)=1/\tau_b+1/[\tau_0+\tau_r\tau_0NSv(1-y)]. \tag{31}$$

This result justifies the procedure utilized in separating τ_b from the multiple trapping effects, e.g., Fig. 11, since the decay rate $1/\tau_b$ adds directly to the old rate defined by Eq. (13).

[24] We have evidence for this latter case in n-type silicon at 300°K.

Recombination Processes in Insulators and Semiconductors

ALBERT ROSE
RCA Laboratory, Princeton, New Jersey
(Received August 24, 1954)

The discrete states in the forbidden zone are divided into ground states and shallow trapping states. The major recombination traffic passes through the ground states. The shallow trapping states cause the observed decay time of free carrier concentrations to exceed the lifetime of a free carrier in the conduction (or valence) band. At low rates of excitation (free carrier concentrations less than ground state concentrations) the electron lifetime and hole lifetime are independent and, in general, significantly different. At high rates of excitation the free electron and hole lifetimes are equal. For an insulator having one class of ground states (a class being defined by the capture cross sections for electrons and holes) the high-light lifetime is bracketed by the two low-light lifetimes.

The behavior of a model having one class of ground states can be described relatively simply and quantitatively. The behavior of a model having more than one class of ground states becomes sufficiently complex that only special cases can be treated easily. More than one class of ground states, however, is required to account for infrared quenching, "superlinearity" and the ability of added ground states to reduce the rate of recombination. These phenomena involve a redistribution of electrons and holes amongst the classes of ground states. Such redistributions can give some meaning to the phrases: "filling of traps" or "saturation of centers."

The recombination behavior of a semiconductor is significantly different from that of an insulator. For example, superlinearity can occur in a semiconductor having only one class of ground states. Also, the photocurrents in a semiconductor can be intrinsically more noisy than the photocurrents in an insulator.

1.0 INTRODUCTION

ONE may increase the density of free electrons and free holes in a semiconductor or insulator by bombarding the material with photons, electrons or other high-energy radiations or by injecting extra carriers as in the transistor. In the steady state the rate of generation of holes and electrons must be equal to their rate of recombination. The way in which these carriers recombine is a primary concern in the relatively separate fields of luminescence, photoconductivity, and semiconductor control devices. While the physics of these recombinations is common to all three fields, the models and modes of description reflect three different approaches. The discrete states in the forbidden zone, through which recombination is most likely to take place, have been called activators and poisoners in luminescence, deathnium and recombination centers in semiconductors, and traps or primary centers in photoconductors. To this list of names can be added donors and acceptors; electron traps and hole traps; deep traps and shallow traps; and an extensive list of centers in the alkali halides.

All of the states just enumerated have the common characteristic of being discrete states in the forbidden zone. Beyond this their chief distinctions and their chief contributions to recombination are described by their capture cross sections for free electrons and for free holes. One might expect, accordingly, to achieve some condensation of terminology. This is one of the aims of the present discussion.

A number of other questions are considered. Perhpas the most important is the question of how much can be said simply, quantitatively and unambiguously about recombination models. It is by no means certain that a simple description of recombination is at all possible. In fact, the present discussion concludes that when there is more than one class of recombination centers, the variety of possible behavior allows only certain special cases to be discussed with reasonable clarity. This variety is inescapable since experience with phosphors, photoconductors, and some of the more insulating semiconductors has already defined broad limits of actually realized behavior.

It is appropriate in a generalized treatment of recombination to try to give a more precise meaning to certain words and phrases that are commonly used but not always accurately defined. The word "trap" is one example. The distinction between "shallow traps" and "deep traps" is a second. And the phrases "filling of traps" and "saturation of recombination centers" are further examples of ideas that at least merit re-examination.

Finally, there are three first order observations which on the surface are puzzling and, indeed, are not readily explained. These observations are (1) photocurrent *versus* light curves that have a power greater than unity and are therefore called "superlinear," (2) the enhancement of photoconductivity by the *addition* of recombination centers, and (3) "infrared quenching" or the attenuation of short-wavelength excitation by the addition of long-wavelength excitation.

The first item is particularly interesting. Usually, one can find a number of models to explain a given observation and one tries to eliminate models. Here it is difficult to find even one model to account for superlinearity. The second item is puzzling because one would expect that the addition of more paths for recombination would shorten rather than lengthen the lifetime of a free carrier. The last item has had frequent discussion in the literature, mainly as infrared quenching of luminescence, but more recently as infrared quenching of photoconductivity as well. All three items, to anticipate later discussion, have this in common: a satisfactory model must contain more than one class of recombina-

tion centers.[1] A "class" is characterized by its capture cross sections for electrons and holes and is independent of energy.

2.0 LIMITATIONS OF THE PRESENT DISCUSSION

Capture is assumed to take place between free carriers and discrete states in the forbidden zone in essentially one step. That is, if capture is a two step process, both steps are assumed to be of negligible duration. The physical merits of this assumption are not argued here. Little enough is known of capture processes. For lack of knowledge, the simpler process is assumed and its possibilities explored.

The discrete states in the forbidden zone are assumed to be physically separate both in the sense that they do not interact and in the sense that electrons and holes cannot make transitions directly between discrete states. At sufficiently high densities such interaction and such transitions undoubtedly do take place. State densities below $10^{17}/cm^3$ are likely to satisfy the criteria of separateness.

The excitation or generation process is assumed to be one in which electrons are raised from the filled band to the conduction band. Free electrons and free holes are thus generated at identical rates. They recombine at the same rate via discrete states in the forbidden zone. Excitation from discrete states into the conduction band is known to occur in many photoconductors and is by no means a negligible process. Many of the ideas of the present discussion can be carried over to aid in the understanding of these more complex processes.[2]

A state in the forbidden zone is assumed to have either one electron or one hole in it. The possibility of its having more than one electron or hole is excluded. This exclusion, again, is not for lack of physical evidence but only to explore simpler assumptions first.

With all of the above simplifications, the problem of recombination still retains a large measure of complexity. There is likely more need for points of view that allow semiquantitative judgments than there is for complete and closed mathematical solutions.

The properties of a single class of discrete states will be discussed first since reasonably simple and quantitative statements can be made about the possible recombination processes. A second class of discrete states will then be introduced to account, at least qualitatively for a number of observations that cannot be accounted for by a single class of states.

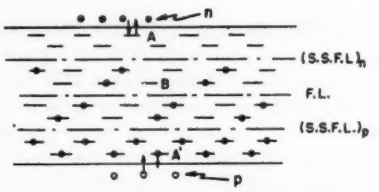


FIG. 1. Model showing that discrete states can be separated for statistical resons into shallow trapping states (A,A') and ground states (B). The ground states are often called recombination centers. $(S.S.F.L.)_s$ is the steady-state Fermi level for electrons.

3.0 DEFINITION OF TERMS

Consider the problem shown in Fig. 1. Free electrons and holes are generated at identical rates and recombine via the discrete states. It is clear, from physical grounds, that there are states (A) near the conduction band such that electrons falling into these states are rapidly re-excited thermally into the conduction band. These states are in thermal equilibrium with the electrons in the conduction band. They will be called shallow trapping states. It is clear also that electrons falling into deeper lying states (B) will not be thermally re-excited for a long time. Before thermal re-excitation takes place, such an electron is more likely to capture a free hole. These deeper lying states will be called ground states. The connotation is that electrons and holes falling into these states have completed their life history. The occupancy of ground states by electrons or holes is determined by the purely kinetic processes of recombination. The occupancy of shallow trapping states is determined by the condition of being in thermal equilibrium with the electrons in the conduction band or the holes in the filled band.

While it is clear that there should be shallow trapping states and ground states, it is not immediately clear what the demarcation line should be. If the discrete states are sufficiently uniformly distributed in energy, and this means that their variation with energy is not as rapid as the function $e^{-E/kT}$, one can locate the demarcation line for electrons at that energy level at which an electron is equally likely to be thermally excited into the conduction band and to capture free electrons. Similarly, a demarcation line for holes can be located at the energy level at which a trapped hole is equally likely to be thermally excited into the conduction band and to capture a free electron. Discrete states lying between these two demarcation lines control the rate of recombination. Discrete states lying outside these two demarcation lines have a finite but negligible effect on recombination. Statistically their contribution to recombination decreases as $e^{-E/kT}$, where E is their energetic distance from the demarcation line.

[1] The problem of "superlinearity" in the field of luminescence has been reported and discussed in a number of papers. All of the papers make use of more than one class of ground states. Recent data are to be found in Nail, Urbach, and Pearlman, J. Opt. Soc. Am. **39**, 690 (1949). Analyses that lead to powers higher than 2 are to be found in S. Roberts and F. E. Williams, J. Opt. Soc. Am. **40**, 516 (1950) and in M. Schön, Z. Naturforsch. **6a**, 251 (1951). More extensive analyses are contained in M. Schön, Tech.-wiss. Abhandl. Osram-Ges. **6**, 49 (1953), and in C. A. Duboc, Brit. J. Appl. Phys. (to be published). The writer is indebted to Dr. Schön for a reprint and to Dr. F. Urbach for a preprint of the last two papers. Insofar as one can characterize analyses made with different types of approximation, the writer believes that the two papers just cited and the present paper are in substantial agreement in using the same physical processes to account for "superlinearity."

[2] Reference 4 is essentially a discussion based on excitation from discrete states.

If there is only one class of discrete states and if the numbers of electrons and holes in the ground states are equal, the demarcation lines coincide with the steady-state Fermi levels[3] for free electrons and holes. This still allows the capture cross section of the ground states for free electrons to be different from their capture cross section for free holes. If the numbers of electrons and holes in the ground states are not equal, both demarcation lines are rigidly shifted by kT times the natural logarithm of the ratio of electrons to holes. Since this shift is usually small (a ratio of 100 to 1 at room temperature causes a shift of only 0.1 electron volt) and since the steady-state Fermi levels have an important connotative value, the steady-state Fermi levels are used in this discussion as a sufficiently good and useful approximation to the demarcation lines.

It there is more than one class of discrete states, each class has its own set of demarcation lines shifted from the steady-state Fermi levels by kT times the natural logarithm of the ratio of electrons to holes in its own ground states.

The ground states, then, are mainly those discrete states that lie between the two steady-state Fermi levels defined by the free electron concentration and the free hole concentration. In this sense, the distinction between traps and ground states is purely statistical. As the excitation is increased and the free carriers increase, the two steady-state Fermi levels move apart toward their respective band edges. States that were trapping states now become part of the group of ground states. Conversely, if the excitation is suddenly removed the free carrier densities decay in time towards zero. As the carrier densities decrease, their Fermi levels recede from the band edges and approach each other at the dark or thermal equilibrium value of the Fermi level. States that were ground states during steady excitation now become trapping states and empty their electrons (or holes) via the free bands. The thermal emptying of these deeper-lying states may be regarded as the emptying of deep traps to be contrasted with the emptying of shallow traps at the start of the decay curve.

It is worth restating that the trap character of a discrete state is not derived here from a physical one-sideness such that electrons can enter and leave through the top but cannot leave through the bottom. The trap character is a purely statistical condition such that an electron is more likely to have an opportunity of leaving through the top by thermal excitation than through the bottom by recombination with a hole.

Over and above the statistical basis for regarding some discrete states as traps, there is evidence that discrete states may be traps in the absolute sense of having a zero or near-zero capture cross section for one sign of carrier.[4] The asymmetry of capture cross sections is discussed in the following section.

4.0 CONCERNING CAPTURE CROSS SECTIONS

The capture cross section of a discrete state for a free carrier can range from about 10^{-13} cm^2 to arbitrarily small values. 10^{-13} cm^2 is the largest value that can be reasonably expected, based on the attractive Coulomb field of a charged center. 10^{-15} cm^2 is a frequently observed value and coincides with the dimensions of a lattice site. Arbitrarily small values can arise from selection rules and from potential barriers surrounding the center—that is, a repulsive Coulomb field. Capture cross sections observed in photoconductors have been reported in the range of 10^{-13} cm^2 to about 10^{-22} cm^2 (reference 7).

Any given center is likely to have markedly different capture cross sections for electrons and holes. For example, a singly negatively charged center is likely to have a large cross section for capturing a hole. After capturing a hole it becomes neutral and is likely to have a much smaller cross section for capturing an electron. A doubly negatively charged center should show an even greater disparity in capture cross section since, after capturing a hole, it actually presents a repulsive field at large distances, for capturing an electron. In the case of CdS photoconducting crystals capture cross sections for electrons have been reported as small as 10^{-21} cm^2.[5,6] On the other hand, attempts to detect the motion of free holes suggest that these same centers have cross sections of 10^{-15} to 10^{-13} cm^2 for holes.

5.0 A SINGLE CLASS OF DISCRETE STATES

This section is a generalization and in some ways a simplification of an earlier discussion[7] in which the emphasis was placed on one sign of carrier. Some of the earlier results will be referred to where there is no need for reporducing the arguments.

5.1 CARRIER CONCENTRATION SMALL COMPARED WITH GROUND STATE CONCENTRATIONS

Figure 2(a) shows an insulator in the dark. A set of discrete states has been assumed lying near the Fermi level. Those below the Fermi level are essentially filled with electrons and those above are essentially empty.

[3] The concentration of free electrons is, in the case of thermal equilibrium, given by $n_e = N_c e^{-E_f/kT}$, where E_f is the energetic distance of the Fermi level from the conduction band. This same relation may be used to *define* a steady-state value of the Fermi level when the steady-state concentration of free electrons is given. The steady-state Fermi level was used in reference 7 and has been used in the analysis of transistors by W. Shockley, *Electrons and Holes in Semiconductors* (D. Van Nostrand Company, Inc., New York, 1950). In the analysis of transistors, it is called a quasi-Fermi level and is a mathematical device to describe the concentration of free carriers. Such use needs no justification other than convenience. The validity of its use in the present paper, however, depends on how faithfully it describes the occupancy of trapping states as well as the occupancy of free states from which it was defined.

[4] J. R. Haynes and J. A. Hornbeck, Phys. Rev. **90**, 152 (1953).

[5] R. W. Smith, RCA Rev. **12**, 350 (1951).

[6] R. H. Bube and S. M. Thomsen, J. Chem. Phys. (to be published).

[7] A. Rose, RCA Rev. **12**, 362 (1951).

In Fig. 2(b) an excitation process has been introduced such that the discrete states lie between the two steady-state Fermi levels and therefore qualify as ground states. The excitation is low enough that the free carrier concentrations, n and p, are kept small compared with the ground state concentrations, n_0 and p_0. This is the usual condition under which photoconductors operate.

If the rate of excitation is $f(\text{cm}^{-3}\ \text{sec}^{-1})$,

$$f = vs_n n p_0 = vs_p p n_0, \tag{1}$$

where v is the thermal velocity of a carrier and s_n, s_p are the capture cross sections of ground states for electrons and holes respectively. In words, this says that, in the steady state, electrons must pour out of the conduction band at the same rate as they are poured in; holes must pour out of the filled band at the same rate as they are poured in, and electrons and holes must pour into the ground states at identical rates. All of the above must be true, else charge would accumulate indefinitely in any of the three places.

The important simplifying fact is this: if the concentrations of free carriers are small compared with the concentrations in the ground states, then, to the same approximation the concentrations of electrons and holes in the ground states remain the same after excitation as they were before excitation. The validity of this statement follows readily from the condition of charge neutrality.

Thus, in Eq. (1), n_0 and p_0 are the concentrations of electrons and holes in the ground states that existed in the dark *before* exposure to light. From Eq. (1):

$$n = f/vs_n p_0, \tag{2a}$$

$$p = f/vs_p n_0. \tag{2b}$$

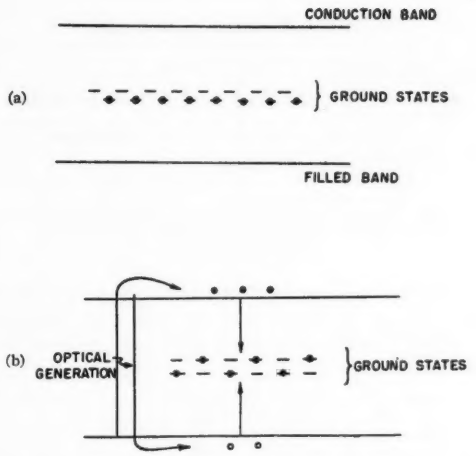
The concentrations of free electrons and holes are determined *independently* by the concentrations and capture cross sections of their respective ground states. From Eq. (2):

$$\text{lifetime of an electron} = (vs_n p_0)^{-1}, \tag{3a}$$

$$\text{lifetime of a hole} = (vs_p n_0)^{-1}. \tag{3b}$$

Because these lifetimes are constants, independent of the number of free electrons, the concentration of free electrons will increase linearly with light intensity or other excitation means. The same is true independently for the concentration of free holes.

When the light is suddenly turned off, the free electron concentration will decay exponentially in time, the concentration falling to e^{-1} of its initial value in $(vs_n p_0)^{-1}$ seconds. The free hole concentration will also decay exponentially and independently at the rate of $(vs_p n_0)^{-1}$ seconds to fall to e^{-1} of its initial value. It is significant here that there are two independent time constants to be associated with a free pair. It is also significant, and to be contrasted with later discussion,



FIGS. 2(a) and (b). Model showing one class of ground states at low excitation rate.

that the lifetime of a free carrier and the observed decay time of its current are one and the same time.

While the numbers of electrons and holes in the ground states remain the same after excitation as before, their distribution does not. Before excitation, the electrons lay below the dark value of the Fermi level and the holes above. After excitation, electrons and holes are uniformly dispersed among the ground states. Any given ground state is occupied on the average $n_0/(n_0+p_0)$ of the time by an electron and $p_0/(n_0+p_0)$ of the time by a hole.

This uniform dispersion of electrons and holes in the ground states can lead to a long-time, low-level tail on the decay curve of photocurrent during which time the thermal distribution of Fig. 2(a) is being re-established. The uniform dispersion also gives some meaning to the phrase "filling of traps." In this case, the upper states that were empty before excitation are now half filled with electrons. Similarly, the lower states that were filled with electrons before excitation, are now only half filled. If the optical excitation had been carried out at a low temperature, this non-equilibrium distribution of electrons and holes would have been "frozen in." Raising the temperature would allow thermal excitation into the free bands and subsequent recombination into the thermal equilibrium distribution of Fig. 2(a). The thermal excitation into the free bands is the seat of the conductivity "glow curve"; the subsequent recombination if radiative is the seat of the luminescence "glow curve."

5.2 BOTH CARRIER CONCENTRATIONS LARGE COMPARED WITH THE GROUND STATE CONCENTRATIONS (FIG. 3)

Let the concentrations of free carriers both be large compared with the concentration of ground states. Then, according to the charge neutrality condition, the

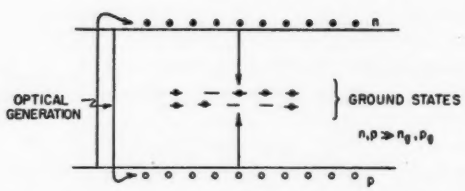


FIG. 3. Model showing one class of ground states at high excitation rate.

concentration of free electrons must be equal to the concentration of free holes—to the approximation given by the ratio of ground states to free carriers. The concentrations of electrons and holes in the ground states need not remain the same after excitation as they were before excitation. Only the sum of the two concentrations is fixed and must, of course, equal the total concentration of ground states. The average occupancy of the ground states by electrons and holes may easily be computed and may bear almost no relation to their initial occupancy.

The rate at which electrons pour into the ground states must equal the rate at which holes pour into the ground states. For this reason and because the concentrations of free electrons and holes are equal, their lifetimes must be equal. (In the previous case of low carrier concentration, the number and lifetimes were, in general, not the same for electrons and holes.) The lifetime is given by:

$$\text{lifetime of electron} = (vs_n p_g)^{-1} = \text{lifetime of hole} = (vs_p n_g)^{-1}. \quad (4)$$

From Eq. (4),

$$n_g = \frac{s_n}{s_n+s_p} N_g, \quad (5a)$$

$$p_g = \frac{s_p}{s_n+s_p} N_g, \quad (5b)$$

where N_g is the total concentration of ground states, n_g+p_g. From Eqs. (4) and (5),

$$\text{lifetime of electron (or hole)} = \left[v\left(\frac{s_n s_p}{s_n+s_p}\right) N_g \right]^{-1}. \quad (6)$$

Equations (5) say that a given ground state will be occupied $s_n/(s_n+s_p)$ of the time by an electron and $s_p/(s_n+s_p)$ of the time by a hole. These are to be contrasted with the corresponding fractions n_g^0/N_g and p_g^0/Ng obtained for the case of low carrier concentrations. n_g^0 and p_g^0 are the initial dark concentrations; n_g and p_g are the concentrations under high light conditions. Only by accident will these two sets of fractions have the same numerical values.

The free electron and hole concentrations are given from Eq. (6) by

$$n = p = f\left[v\frac{s_n s_p}{s_n+s_p} N_g \right]^{-1}. \quad (7)$$

These concentrations are, again as in the low-light case, proportional to the rate of excitation f, since the lifetime of Eq. (6) is a constant.

It is worth noting that at high rates of excitation there is no "saturation" of the ground states that might give meaning to the phrase, "saturation of recombination centers," referred to at the beginning of this paper. The rate of excitation may, indeed, be arbitrarily high without in any sense clogging the recombination paths.[8]

A comparison of the lifetimes at low light intensities, Eq. (3), with the single lifetime at high light intensities, Eq. (6), shows that the two low-light lifetimes either straddle the high-light lifetime or are equal to it. The latter is a fortuitous occurrence when the two low-light lifetimes are equal to each other. In general then, there is an intermediate region between low light intensities and high light intensities in which the two low-light concentrations, n and p, converge to their common high-light value (see Fig. 4). A photoconductivity measurement favors the longer of the two lifetimes and will have the shape of the upper curve in Fig. 4.

The particular form of the convergence of the n and p curves in the intermediate region of Fig. 4 depends upon the values of s_n, s_p, n_g, and p_g. An example of particular interest to photoconductors is described in the next section.

5.3 CARRIER CONCENTRATIONS INTERLACED WITH GROUND STATE CONCENTRATIONS (FIG. 5)

It would be fortuitous to find the two ground state concentrations n_g and p_g equal within a factor of two. It is more likely that the ratio of these concentrations is

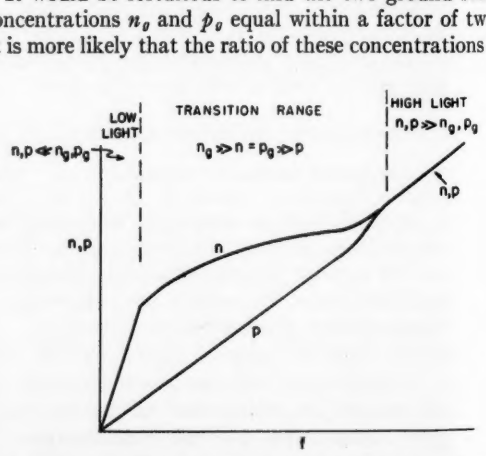


FIG. 4. Schematic description of electron and hole concentrations as a function of rate of excitation. The photocurrents are proportional to these concentrations.

[8] Clogging of the recombination centers could occur if the act of recombination itself occupied a finite time. That time has been assumed to be negligible in this discussion.

measured by factors of ten. This allows the range between high-light intensities and low-light intensities to be described by the conditions:

$$n_0 \gg n \doteq p_0 \gg p, \tag{8}$$

or the parallel set obtained by interchanging n and p. In Eq. (8), n_0 is a constant and equal to its value before excitation. p_0 remains approximately equal to n.

The free electron concentration is, since $p_0 \doteq n$,

$$n = f/vs_n n, \tag{9a}$$

or

$$n = (f/vs_n)^{\frac{1}{2}}. \tag{9b}$$

It is to be noted that this is the first instance of a bimolecular recombination process. The free electron concentration increases only as the square root of the light intensity (Fig. 4). The lifetime of a free electron $(vs_n n)^{-1}$ is, of course, dependent on the free electron concentration and decreases as the electron concentration increases.

The free hole concentration is, by Eq. (8), small compared with the free electron concentration. The free hole concentration, however, increases linearly with light intensity:

$$p = f/vs_p n_0. \tag{10}$$

As the high-light condition is approached, n_0 in Eq. (10) is no longer constant but changes rapidly to its high-light value as shown in Fig. 4.

5.4 EFFECT OF ADDING SHALLOW TRAPPING STATES (FIG. 6)

In the previous sections only ground states and free states were considered. A common characteristic of the recombination processes was that the lifetimes computed for the free carriers were also the observed decay times for photocurrents. The chief effect of the addition of shallow trapping states, as in Fig. 6, is to cause the observed decay time (also rise time) for currents to exceed the lifetime of free carriers by the ratio of trapped to free carriers (see also reference 4). The excitation process must now generate not only the number of electrons in the free states but also those in the shallow trapping states. The rise time in the excitation process is therefore lengthened by the ratio of trapped to free carriers. In the case of some relatively insulating photoconductors, such as are used in the television pickup tubes of the Vidicon[9–11] type, this ratio may easily exceed 10^6. Similarly, when the excitation is removed, not only must the free electrons recombine into ground states but also the shallow trapped electrons must be emptied via the conduction band into the ground states. This lengthens the decay time of the observed photocurrents by the ratio of trapped to free carriers.

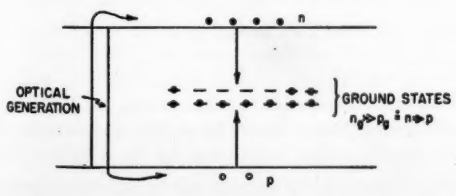


FIG. 5. Model showing conditions for intermediate range of excitations.

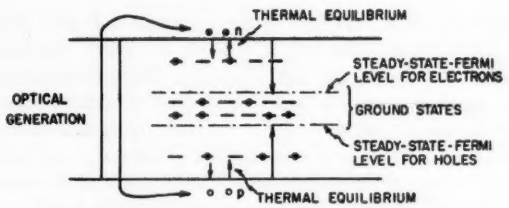


FIG. 6. Model showing both shallow trapping states and ground states.

Since the shallow traps are in thermal equilibrium with the free carriers, the ratio of trapped to free carriers is easily computed to be

$$n_t/n = (N_t/N_c)e^{\Delta E/kT}, \tag{11}$$

where N_t is the number of traps at the level ΔE below the conduction band and $N_c = 10^{19}/\text{cm}^3$ at room temperature. This ratio is not dependent upon the light intensity. Accordingly, the rise and decay times of the photocurrents will be a constant multiple of the free carrier lifetime, and independent of light intensity. It must be remembered that at sufficiently high light intensities the shallow trapping states of Fig. 6 will become part of the group of ground states and the problem reverts then to one of those already discussed in Sec. 5.2.[12]

The introduction of the shallow trapping states also gives some meaning to the phrase "filling of traps." The meaning, however, is perhaps not that normally attached to the phrase. For example, the shallow traps are not actually filled, but only partly occupied in accordance with the Fermi distribution function. Nor does the filling affect the steady-state current but only the rise and decay times. Equations (2) are still valid for the steady-state concentrations.

5.5 EFFECTS OF A CONTINUOUS DISTRIBUTION OF DISCRETE STATES

The models discussed thus far have accounted for linear current-light curves and for observed response times much larger than the actual lifetime of a free carrier in the conduction or filled bands. It is frequently found, however, for the relatively insulating photoconductors in the low-light range, that the photocurrent increases as a fractional power of the light

[9] Weimer, Forgue, and Goodrich, Electronics 23 (1950).
[10] P. K. Weimer and A. D. Cope, RCA Rev. 12, 314 (1951).
[11] Forgue, Goodrich, and Cope, RCA Rev. 12, 335 (1951).
[12] The shallow trapping states of Fig. 6 are assumed to have the same capture cross sections as the ground states.

intensity, the fraction lying between 0.5 and 1.0. Further, the observed speed of response becomes faster with increasing light intensity. The rate of change of speed of response with light intensity is more rapid than can be accounted for by the fractional power of the current-light curve. For example, for CdS crystals the photocurrent has been observed to increase as the 0.9 power of the light intensity, while the decay time for the photocurrent varies almost inversely with light intensity.[5]

Both the fractional powers and the change of speed of response with light intensity are readily accounted for in terms of the model shown in Fig. 7. Here a quasi-continuum of discrete states has been assumed. As the light intensity is increased, the steady-state Fermi levels draw apart and embrace a larger number of ground states. The larger numbers of ground states shorten the lifetimes of the free carriers and thereby lead to fractional powers of the current-light curve. A detailed discussion of various discrete state distributions was carried out in reference 4. There it was shown that if the discrete states are distributed below the conduction band in the form $e^{-\Delta E/kT_c}$, the current light curve should have the power $T_c/(T+T_c)$. Here ΔE is measured from the bottom of the conduction band and T_c is a characteristic temperature greater than or equal to the temperature of measurement T.

Inspection of Fig. 7 shows also how the speed of response can vary rapidly with light intensity. The response time is given by $\tau n_t/u$, where τ is the lifetime of a free electron and n_t/n is the ratio of trapped to free carriers. For a uniform, or near uniform, distribution of discrete states, the lifetime τ is substantially constant. The variation with light intensity just discussed is a slow variation. Also, the number of trapped electrons is substantially constant and given by the number of discrete states within kT of the steady-state Fermi level. The number of free carriers, however, increases almost proportional to the light intensity and reduces the response time at the same rate.

Figure 7 leads also to a t^{-1} form for the decay curve rather than an exponential. This matches frequent observations in which a near linear current-light curve, suggesting a monomolecular process, was accompanied by a t^{-1} decay curve normally characteristic of a bimolecular process.

The model of Fig. 7 is properly subject to the criticism that it is elaborate enough to explain any observations. For this reason two statements need to be emphasized. First, no model short of the complexity of Fig. 7, can logically account for the combined observations of a near linear current-light curve and a response time that varies as the reciprocal light intensity. Second, in spite of the complexity of Fig. 7, there are still several first order observations that cannot be accounted for by this model. These observations require the addition of another class of discrete states. The properties of such a model are discussed in the following section.

6.0 TWO CLASSES OF DISCRETE STATES

Figure 8(a) shows schematically a model with two classes of ground states in thermal equilibrium in the dark. The shallow trapping states are omitted to emphasize the significant properties of this model. Figure 8(b) shows the same model under steady excitation. A particular redistribution of electrons and holes in the ground states has been selected.

At low excitation levels (free carrier concentrations small compared with ground state concentrations) the total numbers of electrons and of holes in the ground states must remain substantially the same after excitation as they were before. A redistribution of these electrons and holes amongst the ground states, however, is free to take place.[13] It is this redistribution that can readily account for (1) activation or sensitization of a photoconductor by the addition of discrete states, (2) infrared quenching, and (3) superlinearity.

The particular redistribution to be expected at low light intensities in the model of Fig. 8 can be obtained in a closed mathematical form. It is an involved expression containing eight parameters—four capture cross sections and four ground state concentrations. The solution for the free electron concentration is reproduced here only to show that this is not a simple problem and to suggest caution in trying to reconstruct the proper model from a few experimental data.

$$n=\frac{f}{v\left[(s_{n1}-s_{n2})\dfrac{G_1}{1+rr_1}+s_{n2}P_0\right]}, \tag{12}$$

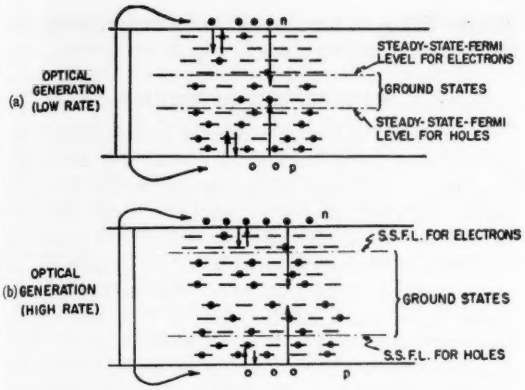
where

$$r=\frac{[N_0(r_1+r_2)-r_1G_1-r_2G_2]\pm\{[N_G(r_1+r_2)-r_1G_1-r_2G_2]^2-4r_1r_2N_0P_0\}^{\frac{1}{2}}}{2r_1r_2P_0},$$

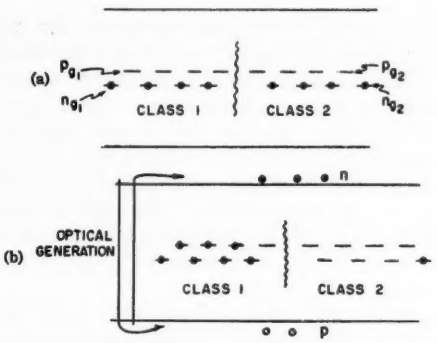
$r_1=s_{n1}/s_{p1}$, $r_2=s_{n2}/s_{p2}$, G_i = sum of ground states in the ith class, N_0 = sum of electrons in both classes, and P_0 the same sum for holes. The physics of the problem is not readily revealed by this solution.

The purpose of the following discussion is to show qualitatively that the three items just enumerated which could not be accounted for by a model having

[13] The redistribution of electrons and holes in the ground states does not take place directly between these states but rather by the intermediate process of excitation into the filled or conduction bands.



FIGS. 7(a) and (b). Model showing continuous distribution of discrete states. Increased rate of excitation adds more states to the ground state category and thereby leads to shorter lifetimes of the free carriers. This model readily accounts for the current-light curves having noninteger powers between 0.5 and 1.0.



FIGS. 8(a) and (b). Model showing two classes of ground states. The classes are drawn separated but are actually physically interspersed. Class 1, by itself, leads to a high rate of recombination. Class 2, by itself, leads to a low rate of recombination.

one class of discrete states can be accounted for by a model having two classes of discrete states.

6.1 ACTIVATION

In Fig. 8 let the discrete states of Class I be such that by themselves the lifetime of free carriers would be short. That is, Class I by itself would make the photoconductor fast and insensitive. The problem is to show that the addition of other discrete states, those in Class II, can actually slow down the recombination process and extend the lifetime of one of the free carriers so that the photoconductor actually becomes more sensitive.

For convenience of discussion, let the four ground state concentrations in Fig. 8a, n_{g1}, p_{g1}, n_{g2}, p_{g2} be nearly equal, n_{g2} exceeding the others by a factor of 2. Let the capture cross sections of the states in Class I be 10^{-15} cm^2 both for free electrons and for free holes. Let the capture cross section of the states in Class II be 10^{-15} cm^2 for free holes and 10^{-20} cm^2 for free electrons. In the absence of the Class II states, the lifetime of a free electron or hole is $(vs_{n1}n_{g1})^{-1}$ with $s_{n1}=10^{-15}$ cm^2. For $n_{g1}=10^{15}$/cm^3, this time is 10^{-7} seconds. The photoconductor is thus fast and insensitive.

The addition of Class II states has the following effect. Holes falling into Class II states become in a sense "trapped." That is, the Class II states do not readily capture electrons. The immediate consequence is that many of the electrons in the Class II states are transferred[13] to the Class I states, thereby sharply reducing the ability of the Class I states to capture free electrons. Free electrons have difficulty in recombining both because the Class II states have a small capture cross section and because the Class I states have a small number of available places. The lifetime of a free electron now approaches $(vs_{n2}n_{g2})^{-1}$, where $s_{n2}=10^{-20}$ cm^2. This time is about 10^{-2} second. The lifetime for free holes, meantime, has been shortened somewhat owing to the increase in the number of hole-capturing ground states. The sensitivity of the photoconductor has been increased by a factor of 10^5 owing to the long-lived electrons. This increase in sensitivity has been brought about by actually adding recombination paths.

The above example is entirely illustrative. There are obviously many choices of parameters that would lead to the same type of result. Also, there are many other behavior patterns derivable from other choices of parameters. The important fact is that the presence of two classes of ground states permits a significant redistribution of electrons and holes to take place. In this example, the second class was so chosen as to sensitize the photoconductor. Other choices can lead to de-sensitization.

Several basic conclusions can be drawn from the above argument. If one calls the Class II states activators or activating centers, these centers need not be hole traps in the absolute sense—that is, having zero capture cross section for electrons. The states need only have a smaller cross section for electrons than for holes. Second, the Class II states are not to be regarded as the "seat" of the photosensitivity. The light does not have to be absorbed by these states nor, in the steady state, does a significant amount of excitation energy have to migrate to the states.[6] The Class II states contribute to sensitivity by effecting a redistribution of electrons and holes among the ground states which redistribution increases the lifetime of all the free electrons. This conclusion is important for a proper estimate of the noise properties of the photoconductor.

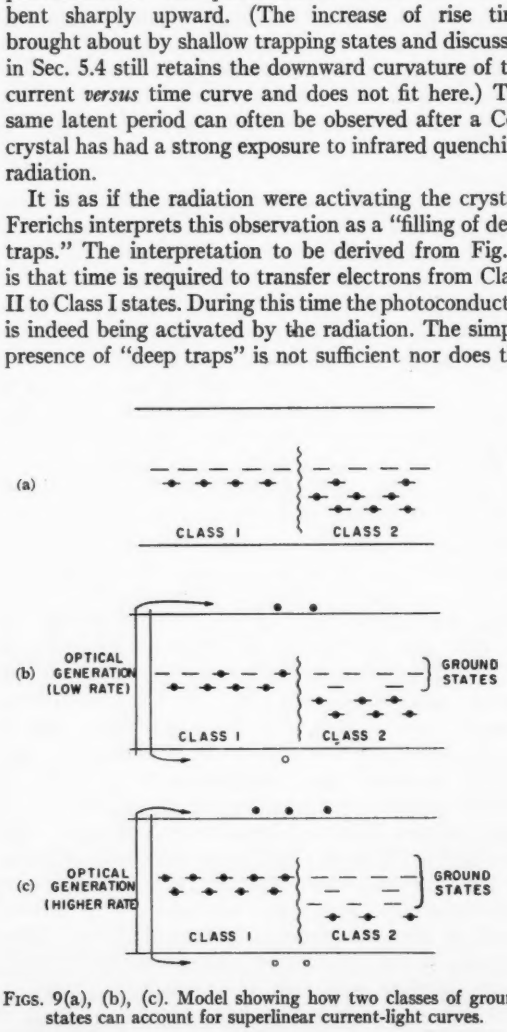
Shulman[14] has shown that the noise properties of a photoconductor with ohmic contacts can be referred back to the noise properties of the incident photon stream. If one argued that activator centers in some way absorbed only part of the incident radiation, the rest of the radiation being somehow "lost" in

[14] Shulman, Smith, and Rose, Phys. Rev. 92, 857 (1953).

insensitive parts of the photoconductor, the photoconductor would be "noisier" than if all of the radiation were effectively used. The present argument leads to the conclusion that activating centers (Class II states) can continuously sensitize the photoconductor by continuously increasing the lifetime of free electrons in such a way that all of the incident photons contribute equally to the photocurrent. The noise properties should then be referred back to all of the incident photon stream even for "partially" activated photoconductors.

The model of Fig. 8 also gives some meaning to the "latent" periods observed by Frerichs[15] and others. Frerichs observed that CdS crystals kept in the dark for a long time and then exposed to a low density of excitation required a long period to develop their photocurrents. The photocurrent *versus* time curve bent sharply upward. (The increase of rise time brought about by shallow trapping states and discussed in Sec. 5.4 still retains the downward curvature of the current *versus* time curve and does not fit here.) The same latent period can often be observed after a CdS crystal has had a strong exposure to infrared quenching radiation.

It is as if the radiation were activating the crystal. Frerichs interprets this observation as a "filling of deep traps." The interpretation to be derived from Fig. 8 is that time is required to transfer electrons from Class II to Class I states. During this time the photoconductor is indeed being activated by the radiation. The simple presence of "deep traps" is not sufficient nor does the phrase "filling of deep traps" completely describe the redistribution taking place in Fig. 8.



FIGS. 9(a), (b), (c). Model showing how two classes of ground states can account for superlinear current-light curves.

[15] R. Frerichs, Phys. Rev. **76**, 869 (1949).

6.2 INFRARED QUENCHING

In terms of Fig. 8(b), the effect of infrared in reducing or quenching the photoconductivity already established by shorter wavelength radiation can readily be understood. Infrared can excite electrons from the filled band into the substantially empty Class II states, the the free holes being captured by Class I states. In this way infrared tends to shift the distribution in the ground states from the sensitive form of Fig. 8(b) to the insensitive form of Fig. 8(a). In brief, its effect on the ground states is exactly counter to the effect of short-wavelength radiation.

After exposure to infrared, a latent period is likely to be encountered during which the short-wavelength radiation re-establishes the favorable distribution in the ground states.

If the Class I states represent radiationless transitions and the Class II states radiation transitions, the infrared quenching model holds also for luminescence.

6.3 SUPERLINEARITY

A model for superlinearity can be derived from Fig. 8 by distributing the Class II states over a range of energies. This is shown in Fig. 9. Figure 9(a) shows the distribution of electrons and holes in the ground states in the dark. Figure 9(b) shows the distribution for a small amount of excitation and Fig. 9(c) for a larger amount of excitation. Essentially, as the excitation is increased the steady-state Fermi levels move apart and include more of the Class II states in the ground state category. This allows steadily more electrons to be transferred from the Class II states to the Class I states. Such transfers tend to sensitize the photoconductor in the manner already described by increasing the lifetime of free electrons. Since the sensitization increases with light intensity the current-light curve is superlinear.

6.4 HIGH EXCITATION

The discussion of the recombination properties of two classes of ground states has thus far been confined to the range of low excitations for which the free carrier concentrations are small compared with the ground state concentrations. Under these conditions the ground states interact with each other in the sense of exchanging their electrons and holes, subject to the condition that the total numbers of electrons and of holes are *separately* maintained constant. At high-light intensities, for which the carrier concentrations are large compared with the ground state concentrations, only the sum total of electrons and holes in each class of ground states remains constant and, of course, equal to the sum of the ground states in that class. At high-light intensities, also, the free carrier concentrations are equal. The consequence is that there

is no longer interaction between the ground states. Each ground state acts independently to adjust its behavior to the free carrier concentration. In particular, each ground state will contain an electron $s_n/(s_n+s_p)$ of the time and a hole $s_p/(s_n+s_p)$ of the time. These occupancy fractions are not affected by the presence of other ground states. A photoconductor cannot be sensitized as it was in the low-light range by the addition of ground states. The addition of ground states must shorten the lifetime of both carriers and desensitize the photoconductor.

The foregoing argument means that the photoconductor of Fig. 8 that was made 10^5 times more sensitive in the low-light range by the addition of Class II states reverts at high light intensities to its initial low sensitivity. The current-light curve will actually appear to saturate at high light intensities because of this low sensitivity. In the intermediate range, the current will have a very low power dependence on light intensity. This power may approach close to zero and give the current-light curve a distinct appearance of saturation. In this intermediate range the free electron concentration is exceeding the concentration of Class II states. In this sense, the Class II states are being saturated. Strictly, however, the saturation of the current-light curve is a reflection of the shift in ground state occupancy from the low-light conditions to the high-light conditions.

7.0 SEMICONDUCTORS

A special problem will be discussed, partly to show how the semiconductor problem differs from the insulator problem and partly to demonstrate the present method of analysis.

Figure 10 shows a semiconductor having one level of ground states N_g below the thermal equilibrium value of the Fermi level.[16] The concentration of thermally produced electrons n is taken to be large compared with the concentration of ground states.[17] The concentration of thermally produced holes is assumed to be negligible. This semiconductor differs from the insulator models already discussed in two significant respects. First, there is a large concentration of free electrons ready to pour into any holes in the ground state. Second, the ground states, prior to generation of additional carriers, are substantially filled with electrons.

7.1 LOW LIGHT EXCITATION

Let the carriers n and p added by excitation be small compared with the ground state concentration

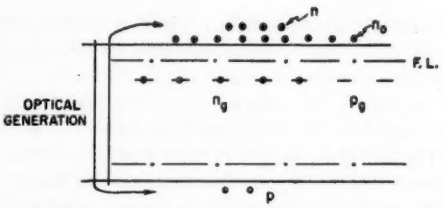



FIG. 10. Semiconductor with one class of ground states. See text for the meaning of the lower dashed line.

N_g. Then

$$f=vs_n(n+n_0)p_g=vs_ppn_g$$
$$\doteq vs_nn_0p_g\doteq vs_ppN_g. \tag{13}$$

From Eq. (13),

$$p/p_g=n_0s_n/N_gs_p. \tag{14}$$

Since s_n and s_p may be chosen arbitrarily, this ratio may be greater or less than unity. If $p\gg p_g$, then $n\doteq p$ and

$$\text{lifetime of electron}=\text{lifetime of a hole}$$
$$=(vs_pN_g)^{-1}. \tag{15}$$

If $p_g\gg p$ (this requires $s_p/s_n\gg n_0/N_g$), then since at all times $n=p+p_g$, it follows that

$$n/p\doteq p_g/p. \tag{16}$$

Equation (16) says that the lifetime of the added electrons will be p_g/p times the lifetime of the added holes.[18] From Eq. (13),

$$\text{lifetime of hole}=(vs_pN_g)^{-1}. \tag{17}$$

From Eqs. (17), (16), and (14),

$$\text{lifetime of electron}=(vs_nn_0)^{-1}. \tag{18}$$

Here, then, the electron and hole concentrations are independently determined by the parameters in Eqs. (17) and (18).

7.2 HIGH-LIGHT EXCITATION

Let the extra carrier concentrations n and p be large compared with the ground state concentration N_g and the thermally produced electrons n_0. This immediately duplicates the high-light insulator problem, namely $n=p$ and

$$\text{electron lifetime}=\text{hole lifetime}=\left(v\frac{s_ns_p}{s_n+s_p}N_g\right)^{-1}. \tag{19}$$

A comparison of the high-light lifetime [Eq. (19)] with the possible low-light lifetimes [Eqs. (15), (17), (18)] shows that the high-light value is always greater than or equal to the low-light values. That is, the semiconductor can be more sensitive at high-light intensities than at low-light intensities. This is to be contrasted

[16] The problem of recombination in a semiconductor having one level of discrete states has been treated more extensively by R. N. Hall, Phys. Rev. 87, 387 (1952) and W. Shockley and W. T. Read, Jr., Phys. Rev. 87, 835 (1952).

[17] Under these conditions, the demarcation line for holes separating shallow trapping states from ground states is located the same distance from the filled band as the dark Fermi level is from the conduction band plus the correction term $kT\times\ln(s_p/s_n)$.

[18] This case is apparently not treated in reference 16. It would account for the fact that the photoconductive decay time does not always give the lifetime of a free pair. The photoconductive measurement gives the longer of the two lifetimes; the lifetime of a free pair is by definition the shorter of the two lifetimes.

with the insulator problem, where a single class of recombination centers could not lead to superlinearity. Two classes were required.

If, in Fig. 10, the single level of discrete states had been located above the Fermi level instead of below it, Eq. (19) for the high-light lifetime would not be altered. However, the low-light lifetime could be made arbitrarily long as the level of discrete states is raised above the Fermi level. The effectiveness of these states for recombination decreases at room temperature by about a factor of 100 for every tenth of a volt above the Fermi level owing to the low thermal density of electrons in them. This is very much the same argument already used in this paper to separate discrete states into shallow trapping states and ground states. Those states lying outside the Fermi limits lose their effectiveness for recombination at the same rate cited above.

If the level of discrete states had been located above the Fermi level, it would then be possible for the high-light lifetimes to be smaller than the low-light lifetimes. Other factors tending in the same direction are discussed in the next section.

7.3 MORE THAN ONE CLASS OF GROUND STATES

If a second class of ground states is added to the semiconductor, its effect in the high-light range (Sec. 7.2) can only be that of adding more recombination paths and of *shortening* the lifetime of both carriers. This statement is independent of the capture cross sections and energy distribution of the Class II states.

The addition of a second class of ground states in the low-light range (Sec. 7.1) can have more varied effects. If the free electron and hole concentrations remain equal after the addition of the Class II states, the lifetime of both carriers is shortened just as it is for the high-light case. This follows from the fact that the addition of more states must shorten the lifetime of a hole, since they provide more places into which a hole can jump. The lifetime of the electron, being equal to that of a hole, must likewise be reduced.

If, in the low-light range, the free electrons and free hole concentrations are not equal, the addition of Class II states must still shorten the lifetime of the minority carrier, but it may either increase or decrease the lifetime of the majority carrier. For example, let an n-type semiconductor have most of its optically created holes in ground states. Under these conditions the number of optically created electrons is essentially equal to the number of holes in ground states. (The conditions of Fig. 10 are assumed here, that is, negligible trapping states between the Fermi level and the conduction band.) There is now a one-to-one correspondence between extra electrons in the conduction band and holes in the ground states. This was not true for the insulator problem discussed in Sec. 5. The addition of Class II states means that some of the holes will be diverted from Class I states to Class II states. If the Class II states have the same capture cross section for electrons as the Class I states, the electron concentration will not be affected. If this capture cross section is greater, the electron concentration will be reduced. Conversely, and this is the significant possibility, if the Class II states have a very small capture cross section for electrons, a large number of holes will congregate in the Class II states and give rise to a large number of extra free electrons. The effect of these states will have been to sensitize the semiconductor as a photoconductor by increasing the free electron concentration. At the same time the lifetime of free pairs has been decreased, since the lifetime of free holes has been decreased.

The type of sensitization just discussed has a fundamental bearing on the noise properties of a photoconductor. In Sec. 5 it was pointed out that a photoconductive insulator could be continuously sensitized by the addition of Class II states without destroying the ability of the photoconductor to reproduce the noise properties of the incident photon stream. In other words, all incident photons make an equal contribution to the photocurrent. In the present instance it becomes possible for the photoconductor to be noisier than the photon stream owing to the addition of sensitizing Class II states. The reason is that it is now possible for only a small fraction of the incident photon stream to contribute holes to the Class II states. At the same time, since the number of holes in the Class II states and their corresponding free electrons are greater than those for the Class I states the observed current will be chiefly that associated with the Class II states. This current must be noisier than the incident photon stream since only a small fraction of the incident photon stream is used in producing the current. (The term "is noisier than" is used here as a shorthand way of saying, "has a smaller signal-to-noise ratio than.")

The fact that Class II states can sensitize a photoconductor in the low-light range and, at the same time, must desensitize it in the high-light range means that the current-light curve of a sensitive semiconducting photoconductor will appear to saturate at high lights. Effects of this kind have been observed for semiconducting crystals CdS.[19] The term "saturation of recombination centers" takes on here a very simple meaning since the number of holes in the ground states must be essentially equal to the number of added electrons in the conduction band. Accordingly, a given class of ground states can support only as many free electrons as there are states in this class. As the number of added free electrons exceeds the number of states in a given class, that class is saturated and other less sensitive classes are forced to take part in the recombination process.

[19] R. H. Bube and R. W. Smith (unpublished).

SUMMARY

In the terms of the present discussion various recombination problems are identified by choosing one characteristic in each of the following categories:

material:	insulator, semiconductor;
excitation range:	low, intermediate, high;
character of ground states:	one class, more than one class
lifetime:	electron, hole, free pair.

Over thirty problems are defined by this table. It is not appropriate here to try to summarize these problems individually but rather to see what general remarks can be made.

(1) In the "high-light" range, electron, hole, and free pair lifetimes are all equal.

(2) In the "high-light" range, the addition of discrete states has only one effect, namely, to reduce the lifetime of both electrons and holes.

(3) In the low-light range, electron and hole lifetimes are independent, as are their concentrations. This is always true for insulators and almost always true for semiconductors. The exception occurs for semiconductors in which the capture cross sections allow electron and hole concentrations to be equal (Sec. 6.1).

(4) In the low-light range the addition of discrete states may reduce the lifetimes of both electrons and holes or may reduce the lifetime of one carrier and increase the lifetime of the other carrier. The latter possibility constitutes the activation process whereby a photoconductor is made more sensitive.

(5) In the low-light range the addition of discrete states always reduces the lifetime of a free pair, since the lifetime of a free pair takes on the shorter of the two carrier lifetimes.

(6) In the high-light range, the carrier concentrations always increase linearly with increase in rate of excitation.

(7) In the low-light range the power of the carrier concentration *versus* excitation curve may take on any values greater than 0.5. Values between 0.5 and 1.0 are derivable from distributions in energy of a single class of discrete states. Values greater than unity are derivable from a distribution in energy of more than one class of discrete states.

(8) If an insulator or semiconductor has been highly sensitized in the low-light range by the addition of more than one class of discrete states, then (see Fig. 4) the majority carrier concentration *versus* excitation curve is likely to have a power less than or equal to 0.5 in the intermediate range.

(9) In the intermediate range, the power of the minority carrier concentration *versus* excitation curve is likely to be greater than unity.

(10) The capture cross sections of discrete states for electrons and for holes are likely to be markedly different owing to the presence of an attractive Coulomb field for one sign of carrier and a zero or repulsive Coulomb field for the opposite sign of carrier.

(11) An ideal photoconductor is a noiseless transformer of photon current into photocurrent. A photoconducting insulator is more likely to retain this property during activation than is a photoconducting semiconductor.

ACKNOWLEDGMENTS

The writer has discussed many of the aspects of this paper with D. O. North. Some of the conclusions emerged from these joint discussions. In addition the writer is indebted to D. O. North, R. H. Bube, and W. M. Webster, all of this laboratory, for final critical readings of the paper.

Theory of the Faraday and Kerr Effects in Ferromagnetics*

PETROS N. ARGYRES†
Department of Physics, University of California, Berkeley, California
(Received August 20, 1954)

Both the Faraday and (magneto-optic, polar) Kerr effects in ferromagnetics are treated on the basis of the band theory of metals. The spin-orbit interaction gives the electron wave functions such left-right asymmetry that the "magnetic" electrons, under the action of a plane polarized light wave, produce an average current perpendicular to the plane of polarization. The polarizability and conductivity tensors are evaluated. The model is capable of describing the rotation of the plane of polarization of the light and the elliptical polarization resulting from transmission or reflection on the ferromagnetic medium. Order of magnitude estimates of the tensor components, based on plausible assumptions on the nature of the electronic wave functions and the energy bands in ferromagnetics, give values that agree reasonably well with experimental results. The temperature and frequency dependence of these effects as given by the formulas is also in agreement with experiment.

1. INTRODUCTION

IT was observed in the last century that the transmitted and reflected beams, into which a plane polarized beam of light splits upon incidence on a ferromagnetic body magnetized in a direction parallel to the beam, become elliptically polarized with their major axes rotated with respect to the plane of polarization of the incident beam. This phenomenon, when it refers to the transmitted beam, is designated as the Faraday effect; it is known as the (magneto-optic, polar) Kerr effect when it refers to the reflected beam. Experiments have shown that these effects are all proportional to the net magnetization of the sample and not to the external magnetic field, as is the case with the nonferromagnetic bodies. The difference between ferromagnetic and nonferromagnetic materials extends to the order of magnitude of the effects. For example, in quartz the Faraday rotation is of the order of 2° per centimeter in an external magnetic field of the order of 10^4 gauss, whereas for iron it is of the order of 380 000° per centimeter under the same conditions. It has been established that these effects are connected with the ferromagnetic properties of the specimen, since for temperatures higher than the Curie temperature of the material these effects disappear along with the ferromagnetic properties.

Macroscopically all these effects can be described, as we shall see below, by assigning to the medium a "refractive" tensor which takes the place of the ordinary complex index of refraction, or, equivalently, by two different complex indices of refraction for light of right- and left-handed circular polarization.

Early attempts to explain these phenomena on a microscopic theory of matter[1] consisted in using the Becquerel formula for the difference of the indices of refraction for right- and left-circularly polarized light under the action of an external magnetic field, with the assumption that in ferromagnetics there is an effective magnetic field active of considerably higher order of magnitude than that of the external field. In fact, Voigt found that the effective field necessary to produce the observed effects is of the order of 10^6-10^7 oersteds, i.e., of the order of magnitude of the so-called Weiss field, which was postulated to account for the existence of the ferromagnetic properties of matter.

It is impossible to explain the origin of such a strong magnetic field. The effective magnetic field for a charge inside a magnetized medium is, according to Wannier,[2] equal to $\mathbf{H}+2\pi(1+p)\mathbf{M}$, where p is a parameter depending on the motion of the electron, varying between 0 and 1. The nature of the Weiss field, responsible for ferromagnetism, was explained by Heisenberg as a result of exchange interactions among the electrons. Although such an interaction energy can be thought of as an equivalent effective magnetic field as far as the alignment of the elementary magnets is concerned, it cannot affect the motion of the electrons as an equivalent magnetic field. It is the motion of the charges that is of importance here, since this motion gives rise to the electric current and thus affects the optical properties of the specimen.

The answer to this problem was provided by Hulme,[3] who introduced into the picture the spin-orbit interaction. This is the energy of interaction of the magnetic moment of an electron, $\boldsymbol{\mu}$, with the magnetic field it "sees" as it moves through the electric field, $-\nabla V$, inside the medium with momentum, $\mathbf{p}$, and it has the form, $\sim\boldsymbol{\mu}\times\nabla V\cdot\mathbf{p}$. We see immediately that such an interaction provides a relation between the motion ($\mathbf{p}$) and the magnetic moment ($\boldsymbol{\mu}$) of the electron, and thus it is plausible that it may be responsible for the connection between the optical and ferromagnetic properties that the Faraday and Kerr effects indicate. Indeed, the spin-orbit interaction can, in a certain approximation, be thought of as an effective magnetic

* Based on a thesis submitted in partial satisfaction of the requirements for the degree of Doctor of Philosophy at the University of California.

† Now at Westinghouse Research Laboratories, East Pittsburgh, Pennsylvania.

[1] For an account of these attempts see: Handbuch der Experimental-Physik, XVI, 1. Teil.

[2] G. Wannier, Phys. Rev. 72, 304 (1947).

[3] H. R. Hulme, Proc. Roy. Soc. (London) A135, 237 (1932); see also O. Halpern, Ann Physik [5] 12, 181 (1932).

field, of vector potential $\sim\mu\times\nabla V$, on the motion of the electrons. This also indicates why the spin-orbit interaction, although present in all matter, is operative in this manner only in ferromagnetics. In nonferromagnetics the electrons are in pairs with opposite μ, and thus half of the electrons find themselves in one magnetic field and the other half in an opposite field, the net result being no effect at all.

Hulme calculated the two indices of refraction, using the Heisenberg model of a ferromagnetic and the Kramers-Heisenberg dispersion formula, which gives the index of refraction in terms of the energy eigenvalues and the matrix elements of the appropriate electric dipole moment operator with respect to the eigenfunctions of the system. He accounted for the difference between the indices of refraction of the right- and left-circularly polarized beams by considering the splitting of the energy eigenvalues of the system due to the spin-orbit interaction and the fact that now states with different energy differences are combined under the two appropriate dipole moment operators. He neglected, however, the effect of the spin-orbit interaction on the wave functions. Kittel[4] showed, by an order of magnitude argument on a simple atomic model, that this change of the wave functions can give rise, on the basis again of the Kramers-Heisenberg dispersion formula, to a difference of the two indices of refraction of the desired order of magnitude. This is important, since several experiments have shown that the orbital angular momentum in ferromagnetic materials is quenched and thus there is no shifting of the energy eigenvalues resulting from spin-orbit interaction. Also, on account of the adopted model, Hulme found only the real indices of refraction, thus neglecting absorption. However, consideration of absorption is essential in discussing all the effects under study here, as can be seen from Eqs. (38), (39), (40), and (41). In fact, if, e.g., we assume the extinction coefficient for iron equal to zero, we can see that the Faraday rotation changes sign. Similarly, the Faraday elliptical polarization and the Kerr rotation vanish unless there is a difference between the coefficients of extinction of right- and left-circularly polarized light.

In order to describe both dispersion and absorption in a direct manner, the band theory of metals is used, on the basis of which the phenomenon of ferromagnetism has been discussed by Slater,[5] Stoner,[6] and others.[7] The main point here is that the exchange forces effectively displace the energy bands of electrons with spin "up" with respect to those of electrons with spin "down," and thus in equilibrium there are more electrons with one spin than with the opposite spin.

This is the distribution of electrons among the possible Bloch states in a single ferromagnetic domain. It should be noted that the actual ferromagnetic samples consist of many domains with their net spins oriented in different directions and that, in addition, the samples are in general polycrystalline. This structure of the specimen has been taken into account in the calculations below.

Since we are interested in computing only the coherent radiation scattered by the system, which alone determines the refraction and extinction of the incident electromagnetic wave, the problem can be treated on the basis of the semiclassical theory of radiation. According to this theory, the current induced in the system by the incident radiation is considered as giving rise to the coherent scattered radiation according to the laws of classical electrodynamics. Thus, the main task is to calculate the current density induced in the system by an electromagnetic wave of certain frequency and, hence to find the conductivity and polarizability tensors of the system under consideration.

The calculation of the induced current had presented some difficulties which have been resolved by Wilson.[8] He showed that if all frequencies are kept in the expression for the induced current, both the polarization and conduction currents can be obtained directly, and there are no infinities to be explained away. The calculations below are carried out in a manner analogous to that of Wilson.

In the following the first-order effect of the spin-orbit interaction on the optical properties of ferromagnetics is examined on the basis of the band theory of metals, and it is shown that such an effect can account for the order of magnitude, dispersion, and temperature dependence of the Faraday and (magneto-optic, polar) Kerr phenomena in ferromagnetics.

2. CALCULATION OF THE TENSORS

The first step is to take as a system a single crystal of a ferromagnetic substance magnetized spontaneously in a given direction, i.e., to consider a single ferromagnetic domain. It is assumed that the lattice vibrations are of negligible importance in the optical phenomena that are considered. This is justified as long as the period of the incident electromagnetic wave is much smaller than the electron-lattice relaxation time. If τ is this relaxation time and ω the angular frequency of radiation, this condition gives $(1/\omega)\ll\tau$, or $\omega\gg(1/\tau)\sim10^{13}$ per second. Thus, the following considerations apply to optical, ultraviolet, and higher frequencies. It will be seen later that other approximations in the calculations will render the theory inapplicable to frequencies much higher than ultraviolet. Also the influence of the external magnetic field is neglected, since, as mentioned in the introduction, it is not sufficient to produce the observed effects. In the following, the effect on the magneto-optic phenomena of the spin-orbit interaction

[4] C. Kittel, Phys. Rev. **83**, 208 (A) (1951); also (private communication).

[5] J. C. Slater, Phys. Rev. **49**, 537, 931 (1936); **52**, 198 (1937); see also J. C. Slater, Solid State and Molecular Theory Group, Technical Report No. 6 (unpublished).

[6] E. C. Stoner, Proc. Roy. Soc. (London) **A165**, 372 (1938); **A169**, 339 (1939).

[7] See several articles in Revs. Modern Phys. **25**, No. 1 (1953).

[8] A. H. Wilson, Proc. Roy Soc. (London) **A151**, 274 (1935).

only is considered. Thus, the results will be strictly valid in the case of a ferromagnetic specimen magnetized in the absence of a magnetic field. Actually, however, they will be very good approximations even in the case of a ferromagnetic under saturation conditions, since magnetic fields of the order of 10^4 gauss produce magneto-optic effects entirely negligible in comparison to those observed. Thus, the role of the external magnetic field reduces to merely magnetizing the sample in a certain direction.

In this calculation the usual one-electron approximation is used in describing the ground and excited states of the system. The one-electron Hamiltonian

$$\mathcal{H}=\mathcal{H}_0+\mathcal{H}'+\mathcal{H}'',$$

where

$$\mathcal{H}_0=(1/2m)\mathbf{p}^2+V(\mathbf{r}),$$

$$\mathcal{H}'=(1/2m^2c^2)[\nabla V(\mathbf{r})\times\mathbf{p}]\cdot\mathbf{s},$$

$$\mathcal{H}''=(e/mc)\mathbf{A}(\mathbf{r},t)\cdot\mathbf{p}.$$

In these expressions $V(\mathbf{r})$ is the potential energy of an electron in the crystal in the absence of radiation, and represents the averaged influence of the nuclei and all the other electrons on the electron under consideration. $\mathbf{A}(\mathbf{r},t)$ is the vector potential of the electromagnetic field inside the material. Such a potential of a monochromatic light wave of angular frequency $\omega(=2\pi\nu)$ is

$$\mathbf{A}(\mathbf{r},t)=\mathbf{a}(\mathbf{r})e^{-i\omega t}+\mathbf{a}^*(\mathbf{r})e^{i\omega t}. \tag{1}$$

$\mathbf{p}$ is the momentum operator $(\hbar/i)\nabla$, and $\mathbf{s}=(\hbar/2)\boldsymbol{\sigma}$ is the electron spin operator. Finally, m is the mass of the electron and $(-e)$ its charge. $\mathcal{H}'$ is the spin-orbit interaction energy of the electron, of which the significance in these problems was discussed in the introduction. $\mathcal{H}''$ is the interaction of the electron with the electromagnetic field in the material, only the largest term being kept.

The wave equation to be solved is

$$\mathcal{H}\Psi=i\hbar(\partial\Psi/\partial t).$$

The solution is obtained by the method of the variation of constants, $\mathcal{H}''$ being treated as a perturbation. Let the orthonormalized eigenfunctions of $(\mathcal{H}_0+\mathcal{H}')$ be $\phi_{\lambda,s}(\mathbf{r})\cdot\exp(-i\omega_\lambda{}^s t)$, i.e.,

$$(\mathcal{H}_0+\mathcal{H}')\phi_{\lambda,s}(\mathbf{r})=\epsilon_\lambda{}^s\cdot\phi_{\lambda,s}(\mathbf{r}), \tag{2}$$

$$\hbar\omega_\lambda{}^s=\epsilon_\lambda{}^s. \tag{2a}$$

Here λ denotes all the quantum numbers necessary for the space part of the wave function and $s(=\pm1)$ stands for the spin quantum number. If the electron is in the state $\phi_{\lambda,s}(\mathbf{r})$ at $t=0$, i.e., before the application of the light wave, its wave function at time t is, to the first order in the amplitude of the light wave,

$$\Psi_{\lambda,s}(\mathbf{r},t)=\phi_{\lambda,s}(\mathbf{r})\exp(-i\omega_\lambda{}^s t)+\sum_{\lambda',\sigma}a_{\lambda\lambda'}{}^{s\sigma}(t)\phi_{\lambda',\sigma}(\mathbf{r})\exp(-i\omega_{\lambda'}{}^\sigma t), \tag{3}$$

where

$$a_{\lambda\lambda'}{}^{s\sigma}(t)=\frac{ie}{mc}\Bigg[\frac{\exp[i(\omega_{\lambda'\lambda}{}^{\sigma s}-\omega)t]-1}{\omega_{\lambda'\lambda}{}^{\sigma s}-\omega}\times\int\phi_{\lambda',\sigma}{}^*(\mathbf{r})\mathbf{a}(\mathbf{r})\cdot\nabla\phi_{\lambda,s}(\mathbf{r})d\tau+\frac{\exp[i(\omega_{\lambda'\lambda}{}^{\sigma s}+\omega)t]-1}{\omega_{\lambda'\lambda}{}^{\sigma s}+\omega}\times\int\phi_{\lambda',\sigma}{}^*(\mathbf{r})\mathbf{a}^*(\mathbf{r})\cdot\nabla\phi_{\lambda,s}(\mathbf{r})d\tau\Bigg], \tag{3a}$$

with the notation

$$\omega_{\lambda'\lambda}{}^{\sigma s}=\omega_{\lambda'}{}^\sigma-\omega_\lambda{}^s=(\epsilon_{\lambda'}{}^\sigma-\epsilon_\lambda{}^s)/\hbar;\quad\sigma=\pm1,$$

($d\tau$ will always denote integration over the whole crystal). The summation over λ', σ in this formula extends over all possible states; it will be seen later how other considerations change this.

The current density of an electron in a state $\Psi_{\lambda,s}(\mathbf{r},t)$ and in the presence of a vector potential $\mathbf{A}(\mathbf{r},t)$ is given by the formula (part of the current operator that would give terms proportional to magnetization square has been dropped)

$$\mathbf{j}_{\lambda,s}(\mathbf{r},t)=ie\hbar/2m(\Psi_{\lambda,s}{}^*\nabla\Psi_{\lambda,s}-\Psi_{\lambda,s}\nabla\Psi_{\lambda,s}{}^*)-(e^2/mc)\mathbf{A}\Psi_{\lambda,s}{}^*\Psi_{\lambda,s}. \tag{4}$$

Using Eqs. (3) and (3a), the easily verified relations [see Eq. (1)]

$$\mathbf{a}e^{-i\omega t}=\tfrac{1}{2}\mathbf{A}+\frac{i}{2\omega}\frac{\partial\mathbf{A}}{\partial t},$$

$$\mathbf{a}^*e^{i\omega t}=\tfrac{1}{2}\mathbf{A}-\frac{i}{2\omega}\frac{\partial\mathbf{A}}{\partial t},$$

and neglecting terms of second or higher order in the amplitude of the light wave, after some simplification, we find that Eq. (4) takes the form,

$$\begin{aligned}\mathbf{j}_{\lambda,s}(\mathbf{r},t)=ie\hbar/2m(&\phi_{\lambda,s}{}^*\nabla\phi_{\lambda,s}-\text{c.c.})\\&-\frac{e^2\hbar}{4m^2c}\sum_{\lambda',\sigma}\Big\{f_{\lambda',\lambda}{}^{\sigma s}(\phi_{\lambda,s}{}^*\nabla\phi_{\lambda',\sigma}-\phi_{\lambda',\sigma}\nabla\phi_{\lambda,s}{}^*)\int\phi_{\lambda',\sigma}{}^*\mathbf{A}\cdot\nabla\phi_{\lambda,s}d\tau+\text{c.c.}\Big\}\\&-\frac{ie^2\hbar}{4m^2c}\frac{1}{\omega}\sum_{\lambda',\sigma}\Big\{f_{\lambda',\lambda}{}^{\sigma s}(\phi_{\lambda,s}{}^*\nabla\phi_{\lambda',\sigma}-\phi_{\lambda',\sigma}\nabla\phi_{\lambda,s}{}^*)\int\phi_{\lambda',\sigma}{}^*\frac{\partial\mathbf{A}}{\partial t}\cdot\nabla\phi_{\lambda,s}d\tau-\text{c.c.}\Big\}\\&-\frac{e^2}{mc}\mathbf{A}\phi_{\lambda,s}{}^*\phi_{\lambda,s},\end{aligned} \tag{5}$$

where

$$f_{\lambda'\lambda}{}^{\sigma s}{\pm}=\frac{1-\exp[-i(\omega_{\lambda',\lambda}{}^{\sigma s}-\omega)t]}{\omega_{\lambda',\lambda}{}^{\sigma s}-\omega}$$

$$\pm\frac{1-\exp[-i(\omega_{\lambda',\lambda}{}^{\sigma s}+\omega)t]}{\omega_{\lambda',\lambda}{}^{\sigma s}+\omega}. \quad (5a)$$

$\phi_{\lambda,s}$ can be found by perturbation theory. Let $\psi_n(\mathbf{k},\mathbf{r})$ be the orthonormalized eigenfunctions of $\mathcal{H}_0$ and $E_n(\mathbf{k})$ their eigenvalues, i.e.,

$$\mathcal{H}_0\psi_n(\mathbf{k},\mathbf{r})=E_n(\mathbf{k})\psi_n(\mathbf{k},\mathbf{r}), \quad (6)$$

$$\int\psi_n{}^*(\mathbf{k}',\mathbf{r})\psi_m(\mathbf{k},\mathbf{r})d\tau=\delta_{nm}\delta(\mathbf{k}'-\mathbf{k}). \quad (6a)$$

Here the reduced Brillouin zone scheme is used in enumerating the Bloch functions $\psi_n(\mathbf{k},\mathbf{r})$; n denotes the energy band, and $\mathbf{k}$ is the wave vector. $\psi_n(\mathbf{k},\mathbf{r})\alpha(s)$ can be taken as unperturbed eigenfunctions, where $\alpha(s)$, $s=\pm1$, are the spin eigenfunctions of the Pauli spin operator σ_{z0}, i.e.,

$$\sigma_{z0}\cdot\alpha(\pm1)=\pm\alpha(\pm1). \quad (7)$$

Here z_0 denotes an arbitrary direction of quantization of the spin. In the "molecular field" approximation of the exchange forces the quantities $E_n(\mathbf{k})\pm\delta$ are taken as the energy eigenvalues of the eigenfunctions $\psi_n(\mathbf{k},\mathbf{r})\alpha(\pm1)$. Now, since the perturbation $\mathcal{H}'$ has the periodicity of the lattice, it connects only Bloch states of different bands with the same $\mathbf{k}$. This also proves that the Bloch functions are the correct zero-order wave functions for the perturbation $\mathcal{H}'$ for all $\mathbf{k}$, except possibly for few $\mathbf{k}$'s for which there might be degeneracy of different bands. Such cases can be disregarded, as their effect is certainly negligible.

In expression (5) for the current density, terms up to the order of (spin-orbit energy)$\times$(amplitude of light wave) are kept, and thus, effectively,

$$\phi_{\lambda,\pm1}(\mathbf{r})=[\psi_n(\mathbf{k},\mathbf{r})\pm\chi_n(\mathbf{k},\mathbf{r})]\alpha(\pm1), \quad (8)$$

$$\epsilon_\lambda{}^{\pm1}=E_n(\mathbf{k})\pm\delta, \quad (8a)$$

where

$$\chi_n(\mathbf{k},\mathbf{r})=\sum_{m\neq n}b_{nm}(\mathbf{k})\psi_m(\mathbf{k},\mathbf{r}), \quad (8b)$$

with

$$b_{nm}(\mathbf{k})=\frac{-i\hbar^2/4m^2c^2}{E_n(\mathbf{k})-E_m(\mathbf{k})}\times\int\psi_m{}^*(\mathbf{k})(\nabla V\times\nabla)_{z0}\psi_n(\mathbf{k})d\tau. \quad (8c)$$

The other parts of $\phi_{\lambda,\pm1}$ give terms of order higher than that worked with. The diagonal matrix element of $\mathcal{H}'$ is zero, as can be seen directly by making use of the time and space inversion symmetry properties of the Bloch functions, thus exhibiting the quenching of orbital angular momentum.

To this approximation, Eq. (5) takes the form,

$$\mathbf{j}_{\lambda,\pm1}(\mathbf{r},t)=\frac{ie\hbar}{2m}(\phi_{\lambda,\pm1}{}^*\nabla\phi_{\lambda,\pm1}-\text{c.c.})$$

$$-\frac{e^2\hbar}{4m^2c}\sum_{\lambda'}\{f_{\lambda',\lambda}{}^+[\mathbf{K}_{\lambda',\lambda}(\mathbf{A})\pm\mathbf{L}_{\lambda',\lambda}(\mathbf{A})]+\text{c.c.}\}$$

$$-\frac{ie^2\hbar}{4m^2c}\frac{1}{\omega}\sum_{\lambda'}\{f_{\lambda',\lambda}{}^-[\mathbf{K}_{\lambda',\lambda}(\partial\mathbf{A}/\partial t)\pm\mathbf{L}_{\lambda',\lambda}(\partial\mathbf{A}/\partial t)]-\text{c.c.}\}$$

$$-\frac{e^2}{mc}\mathbf{A}\{\psi_\lambda{}^*\psi_\lambda\pm[\psi_\lambda\chi_\lambda{}^*+\text{c.c.}]\}, \quad (9)$$

where

$$\mathbf{K}_{\lambda',\lambda}(\mathbf{A})=(\psi_\lambda{}^*\nabla\psi_{\lambda'}-\psi_{\lambda'}\nabla\psi_\lambda{}^*)\int\psi_{\lambda'}{}^*\mathbf{A}\cdot\nabla\psi_\lambda d\tau,$$

$$\mathbf{L}_{\lambda',\lambda}(\mathbf{A})=(\psi_\lambda{}^*\nabla\psi_{\lambda'}-\psi_{\lambda'}\nabla\psi_\lambda{}^*)\int(\psi_{\lambda'}{}^*\mathbf{A}\cdot\nabla\chi_\lambda+\chi_{\lambda'}{}^*\mathbf{A}\cdot\nabla\psi_\lambda)d\tau+(\psi_\lambda{}^*\nabla\chi_{\lambda'}+\chi_\lambda{}^*\nabla\psi_{\lambda'}-\psi_{\lambda'}\nabla\chi_\lambda{}^*-\chi_{\lambda'}\nabla\psi_\lambda{}^*)\int\psi_{\lambda'}{}^*\mathbf{A}\cdot\nabla\psi_\lambda d\tau,$$

$$f_{\lambda',\lambda}{}^\pm=\frac{1-\exp[-i(\omega_{\lambda',\lambda}-\omega)t]}{\omega_{\lambda',\lambda}-\omega}\pm\frac{1-\exp[-i(\omega_{\lambda',\lambda}+\omega)t]}{\omega_{\lambda',\lambda}+\omega},$$

$$\omega_{\lambda',\lambda}=\omega_{\lambda'}-\omega_\lambda=(E_{\lambda'}-E_\lambda)/\hbar,$$

λ denotes quantum numbers $(n,\mathbf{k})$, λ' denotes quantum numbers $(m,\mathbf{k}')$, and

$$\sum_{\lambda'}\rightarrow\frac{\Omega}{(2\pi)^3}\sum_m\int_{\text{first B.Z.}}d\mathbf{k}'. \quad (\Omega=\text{volume of specimen}).$$

The next step is to find the total current density. In this Hartree approximation,

$$\mathbf{j}=\mathbf{j}_{+1}+\mathbf{j}_{-1}, \quad (10)$$

where

$$\mathbf{j}_{+1}=\textstyle\sum_\lambda\mathbf{j}_{\lambda,+1},\quad \mathbf{j}_{-1}=\sum_\lambda\mathbf{j}_{\lambda,-1}, \quad (10a)$$

the summations extending over states occupied by electrons with the corresponding spin. This is equivalent to assuming a completely degenerate electron distribution. For a nondegenerate distribution, $\mathbf{j}_{\pm1}$ are obtained from $\mathbf{j}_{\lambda,\pm1}$ through Eqs. (10a) by multiplication with the Fermi distribution function. $\mathbf{j}_{\lambda,\pm1}$ are given by Eq. (9) where, according to the time-dependent perturbation theory, λ' in the summation should go over all possible one-electron states, whether occupied or unoccupied. This is a result of the Hartree approximation, where each electron is assumed to be in a definite state and the Pauli principle is taken into account by having not more than two electrons (and in the case of two, with opposite spins) in the same Bloch state. It

was observed[9] in connection with the optical properties of atoms that agreement with experiment obtains, if only the nonoccupied states in Eq. (9) are summed over. This can be explained theoretically on the basis of the Fock approximation, which takes the exclusion principle into account directly by describing the states of the whole system by determinantal wave functions constructed out of the one-electron wave functions used.[10] Thus, $\mathbf{j}$ is given by Eqs. (10) and (10a), where $\mathbf{j}_{\lambda,\pm 1}$ are given by Eq. (9) with the understanding that now λ' stands for any one of the nonoccupied states, which is denoted symbolically by $\lambda'>\lambda$.

Making use of the relations

$$\psi_n(-\mathbf{k},\mathbf{r})=\psi_n^*(\mathbf{k},\mathbf{r}), \quad E_n(-\mathbf{k})=E_n(\mathbf{k}),$$

$$\chi_n(-\mathbf{k},\mathbf{r})=-\chi_n^*(\mathbf{k},\mathbf{r}),$$

which give

$$f_{mn}^{\pm}(-\mathbf{k}',-\mathbf{k})=f_{mn}^{\pm}(\mathbf{k}',\mathbf{k}), \quad \mathbf{K}_{mn}(-\mathbf{k}',-\mathbf{k})$$

$$=\mathbf{K}_{mn}^*(\mathbf{k}',\mathbf{k}), \quad \mathbf{L}_{mn}(-\mathbf{k}',-\mathbf{k})=-\mathbf{L}_{mn}^*(\mathbf{k}',\mathbf{k}),$$

and noting that if state $(n,\mathbf{k})$ is occupied so is $(n,-\mathbf{k})$, the total current density

$$\mathbf{j}=\mathbf{j}^{(0)}+\mathbf{j}^{(1)}, \tag{11}$$

where

$$\mathbf{j}^{(0)}=-\frac{e^2h}{2m^2c}\sum_{\lambda(+)}\sum_{\lambda'>\lambda}(\Re f_{\lambda',\lambda}^{+})\mathbf{K}_{\lambda',\lambda}(\mathbf{A})$$

$$+\frac{e^2h}{2m^2c}\frac{1}{\omega}\sum_{\lambda(+)}\sum_{\lambda'>\lambda}(\Im f_{\lambda',\lambda}^{-})\mathbf{K}_{\lambda',\lambda}(\partial\mathbf{A}/\partial t)$$

$$-\frac{e^2}{mc}\mathbf{A}\sum_{\lambda(+)}\psi_\lambda^*\psi_\lambda, \tag{11a}$$

and

$$\mathbf{j}^{(1)}=-\frac{ie^2h}{2m^2c}\sum_{\lambda(-)}\sum_{\lambda'>\lambda}(\Im f_{\lambda',\lambda}^{+})\mathbf{L}_{\lambda',\lambda}(\mathbf{A})$$

$$-\frac{ie^2h}{2m^2c}\frac{1}{\omega}\sum_{\lambda(-)}\sum_{\lambda'>\lambda}(\Re f_{\lambda',\lambda}^{-})\mathbf{L}_{\lambda',\lambda}(\partial\mathbf{A}/\partial t)$$

$$-\frac{e^2}{mc}\mathbf{A}\sum_{\lambda(-)}(\psi_\lambda\chi_\lambda^*+\text{c.c.}). \tag{11b}$$

$\sum_{\lambda(-)}$ indicates that the summation is to be taken over the states that are occupied *only* by electrons with spin $\alpha(+1)$, i.e., the electrons responsible for the magnetic properties of the sample, and which shall be called magnetic electrons. $\sum_{\lambda(+)}$ denotes a summation over all states occupied by electrons with spin $\alpha(+1)$ *and* all states occupied by electrons with spin $\alpha(-1)$. $\Re$ and $\Im$ denote the real and imaginary parts of a complex quantity, respectively. The term $\sum_\lambda(\phi_{\lambda,\pm 1}^*\nabla\phi_{\lambda,\pm 1}-\text{c.c.})$ has been dropped, since its macroscopic value is zero. $\mathbf{j}^{(0)}$ is the induced current density in the absence of any spin-orbit interaction. $\mathbf{j}^{(1)}$ arises from spin-orbit interaction; only the magnetic electrons contribute to $\mathbf{j}^{(1)}$, and this points out the connection of the magnetic and optical properties of metals in so far as the spin-orbit interaction is responsible for it.

In this semiclassical theory of radiation, variations in a region small compared to the wavelength of the wave are unimportant, and thus the calculations can be simplified by finding the average current density over a unit cell. This restricts the applicability of the results to light waves of frequencies ω such that $(2\pi c/\omega)\gg a$, where a is the lattice constant. Since $a\sim 10^{-8}$ cm, $\omega\ll 10^{18}$ per second; hence, the results do not apply to x-ray or higher frequencies This and the previous approximation of disregarding collisions of electrons with lattice vibrations limit the range of validity of the formulas pretty well to optical and ultraviolet frequencies. However, as will be shown below, this averaging will give a simplified expression for the total current density, part of which will be proportional to the electric field $\mathbf{E}$ of the light wave (conduction current) and the rest proportional to $\partial\mathbf{E}/\partial t$ (polarization current). This was first demonstrated by Wilson[8] in connection with the optical properties of solids.

A coordinate system x, y, z is taken so that the wave propagates parallel to the z-axis. It is assumed that $\mathbf{A}(\mathbf{r},t)$ depends only on z, i.e., $\mathbf{A}(\mathbf{r},t)=\mathbf{A}(z,t)$, and it will be proved later that this is a self-consistent assumption, i.e., that this wave gives rise to such a current density that Maxwell's equations for this density admit of a solution $\mathbf{A}(z,t)$ depending only on z. The averaging calculations are straightforward generalizations of those by Wilson, and as such they will not be reported here. However, this averaging process neglects the momentum of the light wave in comparison to that of the electrons (which is reasonable in the frequencies used), and gives essentially the selection rule $\mathbf{k}'=\mathbf{k}$ for the nonvanishing of the integral $\int\psi_{\lambda'}^*\mathbf{A}\cdot\nabla\psi_\lambda d\tau=\int\psi_m^*(\mathbf{k}',\mathbf{r})\mathbf{A}\cdot\nabla\psi_n(\mathbf{k},\mathbf{r})d\tau$. The Bloch character of the wave functions is used here, namely,

$$\psi_n(\mathbf{k},\mathbf{r})=(1/N^{\frac{1}{2}})e^{i\mathbf{k}\cdot\mathbf{r}}u_n(\mathbf{k},\mathbf{r}), \tag{12}$$

and

$$\chi_n(\mathbf{k},\mathbf{r})=(1/N^{\frac{1}{2}})e^{i\mathbf{k}\cdot\mathbf{r}}w_n(\mathbf{k},\mathbf{r}), \tag{13}$$

where, as can be seen from Eqs. (8b) and (12),

$$w_n(\mathbf{k},\mathbf{r})=\sum_{m\neq n}b_{nm}(\mathbf{k})u_m(\mathbf{k},\mathbf{r}), \tag{14}$$

$b_{nm}(\mathbf{k})$ being given by Eq. (8c). Here $u_n(\mathbf{k},\mathbf{r})$, and consequently $w_n(\mathbf{k},\mathbf{r})$ also, has the periodicity of the lattice and is normalized over a unit cell, i.e.,

$$\int u_n^*(\mathbf{k},\mathbf{r})u_n(\mathbf{k},\mathbf{r})d\tau_0=1.$$

$d\tau_0$ will denote integration over a unit cell, in contradistinction to $d\tau$ that denotes integration over the whole

[9] H. Kramers and R. de L. Krönig, Z. Physik 48, 174 (1938).

[10] F. Seitz, *Modern Theory of Solids* (McGraw-Hill Book Company, Inc., New York, 1940), Chap. XVII.

crystal. Thus, $\psi_n(\mathbf{k},\mathbf{r})$ is normalized over the whole crystal according to Eq. (6a); N is the number of unit cells in the crystal.

Using the relations giving the vector potential in terms of the electric field $\mathbf{E}(z,t)$, i.e.,

$$\mathbf{E}=-\frac{1}{c}\frac{\partial \mathbf{A}}{\partial t}, \quad \frac{\partial \mathbf{E}}{\partial t}=-\frac{\omega^2}{c}\mathbf{A},$$

the result of this calculation, i.e., the average values of $\mathbf{j}$, $\mathbf{j}^{(0)}$, and $\mathbf{j}^{(1)}$, given by Eqs. (11), (11a), and (11b), respectively, is

$$\mathbf{J}(\mathbf{r},t)=\mathbf{J}^{(0)}(\mathbf{r},t)+\mathbf{J}^{(1)}(\mathbf{r},t), \tag{15}$$

where $\mathbf{J}$, the total macroscopic current density (which, by the nature of the calculation, includes both the conduction and polarization current), and $\mathbf{J}^{(0)}$, $\mathbf{J}^{(1)}$ are given by the relations,

$$\mathbf{J}^{(0)}=\boldsymbol{\sigma}^{(0)}\cdot\mathbf{E}+\boldsymbol{\alpha}^{(0)}\cdot(\partial\mathbf{E}/\partial t), \tag{15a}$$

$$\mathbf{J}^{(1)}=\boldsymbol{\sigma}^{(1)}\cdot\mathbf{E}+\boldsymbol{\alpha}^{(1)}\cdot(\partial\mathbf{E}/\partial t). \tag{15b}$$

The $\boldsymbol{\sigma}$'s and $\boldsymbol{\alpha}$'s are conductivity and polarizability tensors, of which the components are:

$$\sigma_{ij}{}^{(0)}=\frac{e^2\hbar}{8\pi^3m^2}\frac{1}{\omega}\sum_{m>n}\int_V[\mathcal{I}f_{mn}{}^{-}(\mathbf{k})]P_{mn}{}^{ij}(\mathbf{k})d\mathbf{k},$$

$$\alpha_{ij}{}^{(0)}=-\frac{e^2\hbar}{8\pi^3m^2}\frac{1}{\omega^2}\int_V \times\left\{\frac{m}{\hbar}\delta_{ij}-\sum_{m>n}[\mathcal{R}f_{mn}{}^{+}(\mathbf{k})]P_{mn}{}^{ij}(\mathbf{k})\right\}d\mathbf{k},$$

$$\sigma_{ij}{}^{(1)}=\frac{e^2\hbar}{8\pi^3m^2}\frac{1}{\omega}\sum_{m>n}\int_v[\mathcal{R}f_{mn}{}^{-}(\mathbf{k})]Q_{mn}{}^{ij}(\mathbf{k})d\mathbf{k},$$

$$\alpha_{ij}{}^{(1)}=-\frac{e^2\hbar}{8\pi^3m^2}\frac{1}{\omega^2}\sum_{m>n}\int_v[\mathcal{I}f_{mn}{}^{+}(\mathbf{k})]Q_{mn}{}^{ij}(\mathbf{k})d\mathbf{k},$$

with

$$P_{mn}{}^{ij}(\mathbf{k})=\int u_m{}^*(\mathbf{k})\frac{\partial u_n(\mathbf{k})}{\partial x_i}d\tau_0 \times\int u_m(\mathbf{k})\frac{\partial u_n{}^*(\mathbf{k})}{\partial x_j}d\tau_0, \tag{16a}$$

$$Q_{mn}{}^{ij}(\mathbf{k})=\frac{1}{i}\left[\int u_m\frac{\partial u_n{}^*}{\partial x_i}d\tau_0 \times\int\left(u_m{}^*\frac{\partial w_n}{\partial x_j}+w_m{}^*\frac{\partial u_n}{\partial x_j}\right)d\tau_0 +\int u_m{}^*\frac{\partial u_n}{\partial x_j}d\tau_0 \times\int\left(u_m\frac{\partial w_n{}^*}{\partial x_i}+w_m\frac{\partial u_n{}^*}{\partial x_i}\right)d\tau_0\right]. \tag{16b}$$

$$f_{mn}{}^{\pm}(\mathbf{k})=\frac{1-\exp[-i(\omega_{mn}-\omega)t]}{\omega_{mn}-\omega} \pm\frac{1-\exp[-i(\omega_{mn}+\omega)t]}{\omega_{mn}+\omega},$$

$$\omega_{mn}=\frac{E_m(\mathbf{k})-E_n(\mathbf{k})}{\hbar}.$$

$\sum_{m>n}$ extends over all nonoccupied bands. $\int_V$ denotes integration over all occupied states, *once* for *each* spin orientation (this includes effectively a summation over n). $\int_v$ denotes integration over the states occupied by *magnetic* electrons only. Here, for convenience, the x, y, z components of a vector have been denoted by the subscripts i, $j(i,\ j=1,2,3)$. $f^{\pm}$ is a function that often appears in collision and radiation problems and has the property,[11]

$$\lim_{t\to\infty}f_{mn}{}^{\pm}(\mathbf{k})=\left[\frac{\mathcal{P}}{\omega_{mn}-\omega}+i\pi\delta(\omega_{mn}-\omega)\right] \pm\left[\frac{\mathcal{P}}{\omega_{mn}+\omega}+i\pi\delta(\omega_{mn}+\omega)\right],$$

where $\mathcal{P}/x$ is called the principal value of x (it behaves like $1/x$ for $x\neq0$, but it is zero for $x=0$), and $\delta(x)$ is the Dirac delta function. Thus, Eqs. (16), in the limit of large t, take on the form,

$$\sigma_{ij}{}^{(0)}=\frac{e^2\hbar}{8\pi^3m^2}\sum_{m>n}\int_V\frac{\delta(\omega_{mn}-\omega)}{\omega}P_{mn}{}^{ij}(\mathbf{k})d\mathbf{k},$$

$$\alpha_{ij}{}^{(0)}=-\frac{e^2\hbar}{4\pi^3m^2}\int_V\left[\frac{m}{2\hbar\omega^2}\delta_{ij} -\sum_{m>n}\frac{\omega_{mn}(\mathbf{k})}{\omega^2(\omega_{mn}{}^2-\omega^2)}P_{mn}{}^{ij}(\mathbf{k})\right]d\mathbf{k}, \tag{17}$$

$$\sigma_{ij}{}^{(1)}=\frac{e^2\hbar}{4\pi^3m^2}\sum_{m>n}\int_v\frac{1}{\omega_{mn}{}^2-\omega^2}Q_{mn}{}^{ij}(\mathbf{k})d\mathbf{k},$$

$$\alpha_{ij}{}^{(1)}=-\frac{e^2\hbar}{8\pi^3m^2}\sum_{m>n}\int_v\frac{\delta(\omega_{mn}-\omega)}{\omega^2}Q_{mn}{}^{ij}(\mathbf{k})d\mathbf{k}, \tag{18}$$

where now the integral sign means the principal value of the integral. Thus, the theory is free of infinities. Here the term containing $\delta(\omega_{mn}+\omega)$ has been dropped, as it contributes nothing to the integral. Taking t very large does not invalidate the use of time-dependent perturbation theory, since this can always be done for a small enough radiation field.

The tensors now can be simplified by making use of the symmetry properties of the wave functions. The

[11] W. Heitler, *The Quantum Theory of Radiation* (Clarendon Press, London, 1954), third edition.

calculations are not carried out here,[12] but only the general scheme of simplification is indicated. The tensor components of $\sigma^{(0)}$ and $\alpha^{(0)}$ are of the general form,

$$p_{ij}=\int_V g[\omega_{mn}(\mathbf{k})]P_{mn}{}^{ij}(\mathbf{k})d\mathbf{k}, \tag{19}$$

and those of $\sigma^{(1)}$ and $\alpha^{(1)}$ are of the form,

$$q_{ij}=\int_v g[\omega_{mn}(\mathbf{k})]Q_{mn}{}^{ij}(\mathbf{k})d\mathbf{k}, \tag{20}$$

where $g(\omega_{mn})$ is some function of $\omega_{mn}(\mathbf{k})$, and $P_{mn}{}^{ij}$, $Q_{mn}{}^{ij}$ are given by Eqs. (16a) and (16b). The region V (or v) in $\mathbf{k}$-space over which we integrate can be divided into a number of equal regions, all of which can be obtained from a "fundamental" region V_f (or v_f) by application of the symmetry operations, R, of the point symmetry group of the lattice.[13] It is known[14] that for a general $\mathbf{k}$, with a nondegenerate star, the following equations are true:

$$E_m(R\mathbf{k})=E_m(\mathbf{k}); \quad \omega_{mn}(R\mathbf{k})=\omega_{mn}(\mathbf{k}); \tag{21}$$

and

$$\psi_m(R\mathbf{k},\mathbf{r})=\psi_m(\mathbf{k},R^{-1}\mathbf{r}); \quad u_m(R\mathbf{k},\mathbf{r})=u_m(\mathbf{k},R^{-1}\mathbf{r}). \tag{22}$$

Equations (19) and (20) can then be rewritten in the following form:

$$\begin{aligned}
p_{ij}&=\int_V g[\omega_{mn}(\mathbf{k})]P_{mn}{}^{ij}(\mathbf{k})d\mathbf{k}\\
&=\sum_R\int_{(RV_f)} g[\omega_{mn}(\mathbf{k})]P_{mn}{}^{ij}(\mathbf{k})d\mathbf{k}\\
&=\int_{Vf}\sum_R g[\omega_{mn}(R\mathbf{k})]P_{mn}{}^{ij}(R\mathbf{k})d\mathbf{k}\\
&=\int_{Vf} g[\omega_{mn}(\mathbf{k})]\sum_R P_{mn}{}^{ij}(R\mathbf{k})d\mathbf{k},\\
q_{ij}&=\int_{vf} g[\omega_{mn}(\mathbf{k})]\sum_R Q_{mn}{}^{ij}(R\mathbf{k})d\mathbf{k}.
\end{aligned}$$

Within the fundamental regions V_f and v_f the functions $\omega_{mn}(\mathbf{k})$, $P_{mn}{}^{ij}(\mathbf{k})$, and $Q_{mn}{}^{ij}(\mathbf{k})$ depend on the detailed nature of the crystal potential $V(\mathbf{r})$, and no general statement can be made about them; we can hope, then, to detect the vanishing or nonvanishing of p_{ij} and q_{ij}, not through the integration over $\mathbf{k}$, but rather through the vanishing or nonvanishing of the factors $\sum_R P^{ij}(R\mathbf{k})$ and $\sum_R Q^{ij}(R\mathbf{k})$ in the integrand.

[12] For more details and other points see author's thesis, University of California, Berkeley, 1954 (unpublished).
[13] H. Brooks, thesis, Harvard University, 1940 (unpublished).
[14] F. Seitz, Ann. Math. 37, 17 (1936); see also Bouckaert, Smoluchowski, and Wigner, Phys. Rev. 50, 58 (1936).

Using Eq. (22), it can be proved that

$$\int u_m{}^*(R\mathbf{k})\nabla u_n(R\mathbf{k})d\tau_0=\int u_m{}^*(\mathbf{k})(R\nabla)u_n(\mathbf{k})d\tau_0,$$

and thus,

$$\sum_R P_{mn}{}^{ij}(R\mathbf{k})=\sum_R\int u_m{}^*(\mathbf{k})(R\nabla)_i u_n(\mathbf{k})d\tau_0 \times\int u_m{}^*(\mathbf{k})(R\nabla)_j u_n(\mathbf{k})d\tau_0.$$

An analogous equation can be written for $\sum_R Q_{mn}{}^{ij}(R\mathbf{k})$. We are interested in the case of the cubic group 0_h; this is the point symmetry group for cubic lattices, and thus the following apply to the ferromagnetic Fe and Ni crystals. In the special case of the beam propagating along one of the cubic axes of the crystal, $(R\nabla)_i$ can be found easily, and thus some tensor components are proved to be zero. By a tensor transformation the tensors for a general arrangement of the direction of propagation of the beam, the cubic axes, and the spin direction can then be found. Then the average tensors over the different ferromagnetic domains and the crystal orientations can be found, and these are

$$\begin{aligned}
\sigma_{ij}&=\sigma_0\delta_{ij}+\sigma^{(1)}\beta_{ij},\\
\alpha_{ij}&=\alpha_0\delta_{ij}+\alpha^{(1)}\beta_{ij},
\end{aligned} \tag{23}$$

where

$$\delta_{ij}=\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}, \quad \beta_{ij}=\begin{bmatrix}0&-\beta_3&\beta_2\\\beta_3&0&-\beta_1\\-\beta_2&\beta_1&0\end{bmatrix},$$

[$\beta_i (i=1,2,3)$ are the direction cosines of the net spin direction of the specimen with respect to coordinate system attached to the beam propagating along the 3-direction].

$$\sigma_0=\frac{e^2\hbar}{8\pi^2m^2}\frac{1}{3}\sum_{m>n}\int_V\frac{\delta(\omega_{mn}-\omega)}{\omega}\left|\int u_m{}^*\nabla u_n d\tau_0\right|^2 d\mathbf{k},$$

$$\alpha_0=\frac{e^2\hbar}{4\pi^3m^2}\frac{1}{3}\sum_{m>n}\int_V\frac{\omega_{mn}}{\omega^2(\omega_{mn}{}^2-\omega^2)} \times\left|\int u_m{}^*\nabla u_n d\tau_0\right|^2 d\mathbf{k}-\frac{e^2n}{m}\frac{1}{\omega^2}, \tag{24}$$

where n = total number of electrons per unit volume.

$$\sigma^{(1)}=\frac{e^2\hbar}{4\pi^3m^2}\sum_{m>n}\int_v\frac{1}{\omega_{mn}{}^2-\omega^2}Q_{mn}(\mathbf{k})d\mathbf{k},$$

$$\alpha^{(1)}=-\frac{e^2\hbar}{8\pi^2m^2}\sum_{m>n}\int_v\frac{\delta(\omega_{mn}-\omega)}{\omega^2}Q_{mn}(\mathbf{k})d\mathbf{k}, \tag{25}$$

with

$$\varrho_{mn}(\mathbf{k})=2\Re\left[\frac{1}{i}\int u_m\frac{\partial u_n^*}{\partial y}d\tau_0 \times\int\left(u_m^*\frac{\partial w_n}{\partial x}+w_m^*\frac{\partial u_n}{\partial x}\right)d\tau_0\right]. \quad (26)$$

w_n is given by Eq. (14), where now

$$b_{nm}(\mathbf{k})=-\frac{i\hbar/4m^2c^2}{\omega_{mn}(\mathbf{k})}\int\psi_m^*(\nabla V\times\nabla)_z\psi_n d\tau.$$

Thus, the average macroscopic total current density

$$\mathbf{J}_{\text{total}}=\boldsymbol{\sigma}\cdot\mathbf{E}+\boldsymbol{\alpha}\cdot\partial\mathbf{E}/\partial t, \quad (27)$$

where $\boldsymbol{\sigma}$ and $\boldsymbol{\alpha}$ are given by Eqs. (23). This is the basic equation which will give all the optical phenomena when used in conjunction with Maxwell's equations for the electromagnetic field.

3. FARADAY AND KERR EFFECTS

Consider now the case of a ferromagnetic sample (such as iron or nickel) magnetized in the $+z$-direction. In such instance,

$$\beta_1=\beta_2=0, \quad \beta_3\neq 0,$$

and the conductivity and polarizability tensors are

$$\boldsymbol{\sigma}=\begin{pmatrix}\sigma_0 & -\sigma_1 & 0\\ \sigma_1 & \sigma_0 & 0\\ 0 & 0 & \sigma_0\end{pmatrix}, \quad \boldsymbol{\alpha}=\begin{pmatrix}\alpha_0 & -\alpha_1 & 0\\ \alpha_1 & \alpha_0 & 0\\ 0 & 0 & \alpha_0\end{pmatrix}, \quad (28)$$

with

$$\sigma_1=\sigma^{(1)}\beta_3; \quad \alpha_1=\alpha^{(1)}\beta_3. \quad (28a)$$

The tensors have the form required by general symmetry arguments except for the last diagonal element, which does not necessarily have to be equal to the other two. This may arise from magnetostriction or other purely optical effects which have been neglected. However, it will be seen that the magnitude of the third diagonal element does not enter the discussion of the Faraday and Kerr effects.

The relevant Maxwell equations are

$$\nabla\times\mathbf{E}=-\frac{1}{c}\frac{\partial\mathbf{H}}{\partial t},$$

$$\nabla\times\mathbf{H}=\frac{1}{c}\frac{\partial\mathbf{E}}{\partial t}+\frac{4\pi}{c}\mathbf{J}_{\text{total}}.$$

It should be noted here that both the so-called conduction and polarization currents have been included in $\mathbf{J}_{\text{total}}$. Also, it has been assumed that $(\partial\mathbf{M}/\partial t)=0$, where $\mathbf{M}$ is the average magnetization of the sample. This assumption is justified, since the induced magnetization by the $\mathbf{H}$ field of the light wave is zero in our frequency range.

Using expression (27) for $\mathbf{J}_{\text{total}}$,

$$\nabla\times\mathbf{E}=-\frac{1}{c}\frac{\partial\mathbf{H}}{\partial t}, \quad (29)$$

$$\nabla\times\mathbf{H}=\frac{1}{c}\frac{\partial\mathbf{E}}{dt}+\frac{4\pi}{c}\boldsymbol{\alpha}\cdot\frac{\partial\mathbf{E}}{\partial t}+\frac{4\pi}{c}\boldsymbol{\sigma}\cdot\mathbf{E}. \quad (30)$$

Looking for solutions of the form of a monochromatic plane wave of angular frequency ω propagating along the z-axis, we have

$$\mathbf{E}=\mathbf{E}_0e^{i\omega(t-Nz/c)}$$

$$\mathbf{H}=\mathbf{H}_0e^{i\omega(t-Nz/c)}, \quad (31)$$

where $\mathbf{E}_0$, $\mathbf{H}_0$ are constant vectors and N is the complex index of refraction. It is known from the electromagnetic theory of optical phenomena that such solutions exist for suitable N. This proves the self-consistency of the assumption made before, namely, that $\mathbf{E}$ is a function of z and t only. Equations (29) and (30) become, for such time and space dependence of $\mathbf{E}$ and $\mathbf{H}$,

$$N(\mathbf{E}\times\mathbf{k})=-\mathbf{H}, \quad (32)$$

$$N(\mathbf{H}\times\mathbf{k})=\mathfrak{A}\cdot\mathbf{E}, \quad (33)$$

where

$$\mathfrak{A}=1+4\pi\alpha+\frac{4\pi}{i\omega}\sigma=\begin{pmatrix}A_0 & -A_1 & 0\\ A_1 & A_0 & 0\\ 0 & 0 & A_0\end{pmatrix},$$

$$A_0=1+4\pi\alpha_0+\frac{4\pi}{i\omega}\sigma_0,$$

$$A_1=4\pi\alpha_1+\frac{4\pi}{i\omega}\sigma_1, \quad (33a)$$

and $\mathbf{k}$=unit vector in the z-direction. Substituting Eq. (32) for $\mathbf{H}$ in Eq. (33), the fundamental equation for the optics of anisotropic bodies is obtained, namely,

$$\mathfrak{A}\cdot\mathbf{E}=N^2[\mathbf{E}-\mathbf{k}(\mathbf{E}\cdot\mathbf{k})], \quad (34)$$

or, in component form,

$$\begin{aligned}(A_0-N^2)E_x-A_1E_y&=0,\\ A_1E_x+(A_0-N^2)E_y&=0,\\ A_0E_z&=0.\end{aligned} \quad (35)$$

The first two equations for E_x, E_y have a nontrivial solution only if the determinant of their coefficients vanishes, i.e., $(A_0-N^2)^2+A_1^2=0$, and thus N^2 is equal to either

$$N_+^2=A_0-iA_1, \quad (36)$$

or

$$N_-^2=A_0+iA_1. \quad (37)$$

The solution corresponding to $N_+^2=A_0-iA_1$ is $E_y^{(+)}=+iE_x^{(+)}$, and the one corresponding to $N_-^2=A_0+iA_1$

is $E_y^{(-)}=-iE_x^{(-)}$. Thus, there are two possible ways for a wave to propagate along the z-axis: either as $\mathbf{E}_0^{(+)}\exp\{i\omega[t-(N_+/c)z]\}$, which denotes a right-handed circularly polarized wave traveling with the complex velocity c/N_+; or as $\mathbf{E}_0^{(-)}\exp\{i\omega[t-(N_-/c)z]\}$, which is a left-handed circularly polarized wave traveling with complex velocity c/N_-. Any linear superposition of the two is also a possible wave.

The constants pertaining to the Faraday and (magneto-optic, polar) Kerr effects can now be found. A plane polarized beam, upon incidence on the surface of a sample, assuming normal incidence, splits into the refracted and reflected beams. The refracted beam becomes elliptically polarized with the major axis rotated by an angle

$$\phi_F=\tfrac{1}{2}(\omega z/c)\Re(N_+-N_-), \tag{38}$$

at a distance z from the boundary (a positive sign of ϕ_F denotes a rotation of the major axis from x- to y-axis; xyz is a right-handed triad), and with an ellipticity (ratio of the minor to the major axis) given by

$$\epsilon_F=-\tanh[\tfrac{1}{2}(\omega z/c)\Im(N_+-N_-)]. \tag{39}$$

Equations (38) and (39) are the constants of the Faraday effect. The reflected beam also becomes elliptically polarized with the major axis rotated by an angle

$$\phi_K=-\Im[(N_+-N_-)/(N_+N_--1)], \tag{40}$$

with an ellipticity equal to

$$\epsilon_K=-\Re[(N_+-N_-)/(N_+N_--1)]. \tag{41}$$

Equations (40) and (41) are the constants of the (magneto-optic, polar) Kerr effect, in the special case of normal incidence. ϕ_K, too, counts as positive for rotation from x- to y-axis, but it must be kept in mind that the beam to which it refers is traveling along the $-z$-direction. These formulas are valid for $(N_+-N_-)\ll N_+$ or N_-.

It is interesting to note that all four phenomena can be observed only in absorbing media. If the medium is transparent, i.e., with real N_+ and N_-, then the ellipticity of the refracted beam is zero, i.e., it is a plane polarized beam, and the rotation of the polarization of the reflected beam also vanishes.

The phenomena connected with the case of a ferromagnetic magnetized in a direction perpendicular to the propagation of the beam, such as the Cotton-Moutton effect, etc., cannot be discussed, since a consistent treatment of this problem requires the calculation of the tensor components to the second order in spin-orbit interaction.

N_+, N_- are now calculated for our case. From Eqs. (33a) it can be seen that

$$A_0=1+4\pi(\alpha_0-i\sigma_0/\omega), \tag{42}$$

$$iA_1=4\pi[(\sigma_1/\omega)+i\alpha_1]. \tag{43}$$

From the definitions of σ_0, α_0, σ_1, α_1, it is noted that $(\sigma_1/\omega)\ll\alpha_0$ and $\alpha_1\ll(\sigma_0/\omega)$, since only the magnetic electrons contribute to σ_1, α_1, whereas all of them contribute to σ_0, α_0. This is especially so near the Curie point, since at the Curie temperature $\sigma_1=\alpha_1=0$. Experimental values, on the other hand, indicate that at room temperatures (σ_1/ω) and α_1 are about 30–50 times smaller than α_0 and (σ_0/ω), respectively. Using this approximation, $iA_1\ll A_0$, and Eqs. (36) and (37),

$$N_+=(A_0-iA_0)^{\frac{1}{2}}\cong\sqrt{A_0}-\tfrac{1}{2}iA_1/\sqrt{A_0},$$

$$N_-=(A_0+iA_1)^{\frac{1}{2}}\cong\sqrt{A_0}+\tfrac{1}{2}iA_1/\sqrt{A_0}.$$

Thus, in this approximation,

$$N_+-N_-=-iA_1/\sqrt{A_0}, \tag{44}$$

$$\frac{N_+-N_-}{N_+N_--1}=\frac{-iA_1}{(A_0-1)\sqrt{A_0}}. \tag{45}$$

It should be observed that $\sqrt{A_0}$ is the complex index of refraction, $N=n-ik$ (n=real index of refraction; k=extinction coefficient), of the medium in the absence of any spin-orbit interaction. This is approximately equal to the complex index of refraction of the ferromagnetic above the Curie point, as can be seen with the help of Eqs. (24) and verified experimentally by reflectivity *vs* temperature experiments.[15]

Thus, from Eqs. (42), (43), (44), and (45),

$$N_+-N_-=(-4\pi)\frac{(\sigma_1/\omega)+i\alpha_1}{n-ik}, \tag{46}$$

$$\frac{N_+-N_-}{N_+N_--1}=(-4\pi)\frac{(\sigma_1/\omega)+i\alpha_1}{(n-ik)[(n-ik)^2-1]}. \tag{47}$$

These are to be used to estimate the Faraday and Kerr constants given by Eqs. (38), (39), (40), (41).

4. COMPARISON WITH EXPERIMENT

In comparing the results of this theory with experiment, n and k are assumed to be given by independent experiments on the optical constants of the specimen, although they are given in principle through Eqs. (24). We shall attempt to estimate the fundamental quantities of our theory, σ_1 and α_1, in terms of which the Faraday and Kerr effects are given.

The expressions for σ_1 and α_1 [Eqs. (25)] can be approximated as follows:

$$\sigma_1=\frac{e^2h}{4\pi^2m^2}\beta_3\sum_{m>n}\langle Q_{mn}/(\omega_{mn}^2-\omega^2)\rangle_{kv}\int_v d\mathbf{k},$$

$$\alpha_1=-\frac{e^2h}{8\pi^2m^2}\beta_3\frac{1}{\omega^2}\sum_{m>n}\langle\delta(\omega_{mn}-\omega)Q_{mn}\rangle_{kv}\int_v d\mathbf{k},$$

[15] L. Ornstein and J. H. van der Veen, Physica **3**, 289 (1936); L. Ornstein and O. Koefoed, Physica **5**, 175 (1938); J. H. van der Veen and L. Ornstein, Physica **6**, 439 (1939).

where the angular parentheses indicate an obvious average. But now

$$\int_v d\mathbf{k} = (2\pi)^3 \times \left(\frac{\text{no. of magnetic electrons}}{\text{unit volume}}\right) = (2\pi)^3 \frac{M_s}{\frac{1}{2}g\mu_B},$$

where M_s is the saturation magnetization of a domain, g is the spectroscopic splitting factor, and $\mu_B = (e\hbar/2mc)$ is the Bohr magneton. On the other hand, β_3 is the average value of the direction cosine of the net spin direction with respect to the z-axis, and thus, since the magnetic moment of the electron is opposite to its spin, M, the z-component of the net magnetization, is given by $M = -M_s\beta_3$. Combining the last four equations and taking $g \approx 2$,

$$\sigma_1 = -\frac{4ec}{m}\{\sum_{m>n} \langle Q_{mn}/(\omega_{mn}^2 - \omega^2)\rangle_{Av}\} M, \tag{48}$$

$$\alpha_1 = \frac{2\pi ec}{m}\left\{\frac{1}{\omega^2}\sum_{m>n} \langle \delta(\omega_{mn}-\omega) Q_{mn}\rangle_{Av}\right\} M. \tag{49}$$

Since n,k depend only slightly on magnetization, as explained above, it can be seen immediately that the Faraday and Kerr constants are proportional to the net magnetization M of the specimen. ϵ_F is always observed for so small a film width, z, that it is always proportional to $z(N_+ - N_-)$ [see Eq. (39)]. This is what is observed experimentally, as was mentioned in the introduction. (See also footnote 1.)

The proportionality constant depends on frequency, but it is temperature-independent. All temperature dependence of the Faraday and Kerr constants is thus given by $M(T)$ according to our theory. It should be noted that this is so mainly because, in the calculations for optical and ultraviolet frequencies, the effect of electron–phonon collisions was neglected. The effect of temperature on the electron distribution is certainly negligible. The available experimental evidence[1] indicates that this is indeed the case at these frequencies, the magneto-optic effects under consideration here decreasing slowly at first with rising temperature, much faster as the Curie point is approached, and finally vanishing at the Curie point, much in the same way the M *vs* T curve behaves.

The dispersion of the effects is also given in principle by our formulas, but it is rather difficult to state it explicitly, owing to the complexity of the integrals involved and the lack of accurate knowledge of $\omega_{mn}(\mathbf{k})$, $Q_{mn}(\mathbf{k})$. A similar situation exists with the optical constants n and k. However, the fact that σ_1 depends on ω through the expression $(\omega_{mn}^2 - \omega^2)$ gives rise to the possibility for the Faraday rotation to change sign (an experimentally observed fact) below a certain frequency; whereas an effective magnetic field approach to the problem gives a dependence like $(\omega_{mn}^2 - \omega^2)^2$, which excludes such a possibility.

In order to get a numerical estimate of the order of magnitude of these effects, it is necessary to make further approximations as to the structure of the wave functions and the energy bands. We first approximate the integrals in the expression (26) for $Q_{mn}(\mathbf{k})$, the average value of which appears in Eqs. (48) and (49). Introducing the abbreviations

$$\int u_m^* \frac{\partial}{\partial x} u_n d\tau_0 = \langle m|\partial_x|n\rangle,$$

$$-\frac{i\hbar^2}{4m^2c^2}\int \psi_l^*\left(\frac{\partial V}{\partial x}\frac{\partial}{\partial y} - \frac{\partial V}{\partial y}\frac{\partial}{\partial x}\right)\psi_n d\tau = (l|\mathcal{O}|n),$$

we have

$$Q_{mn} = \frac{1}{i}2\langle m|\partial_y|n\rangle^*\left[\sum_{l\neq n}\frac{(l|\mathcal{O}|n)}{E_n - E_l}\langle m|\partial_x|l\rangle + \sum_{l\neq m}\frac{(l|\mathcal{O}|m)^*}{E_m - E_l}\langle l|\partial_x|n\rangle\right].$$

Here $\Re$ has been dropped, since the matrix elements of ∂_x, ∂_y, and $\mathcal{O}$ are purely imaginary numbers, and thus the whole expression is real. As will be seen below, the most important m-band is the 4s-band, whereas n stands for one of the 3d-bands, namely that occupied by most of the magnetic electrons. Thus, in this crude estimate, the second sum compared to the first in the previous expression for Q_{mn} may be neglected, since $E_n - E_l$ can become quite small with l any one of the other overlapping 3d-bands. Furthermore, the terms making the major contribution to the first sum are the ones with l one of the 3d-bands. Therefore, we shall take

$$Q_{mn} = \frac{2}{\Delta E}\frac{1}{i}\langle m|\partial_y|n\rangle^* \sum_l (l|\mathcal{O}|n)\langle m|\partial_x|l\rangle, \tag{50}$$

where

$$\Delta E = \langle E_n(\mathbf{k}) - E_l(\mathbf{k})\rangle_{Av}, \tag{51}$$

and l denotes any one of the 3d-bands not occupied by most of the magnetic electrons (this is the n-band). The Bloch functions for the 3d-bands have been calculated by Fletcher[16] for nickel on the basis of the tight-binding approximation according to which Bloch functions are constructed as linear combinations of atomic wave functions centered about different lattice points. Since the five 3d-bands overlap, a linear combination of all five 3d-atomic wave functions should be taken for the construction of the corresponding Bloch functions with different coefficients for different values of the wave vector $\mathbf{k}$. Even this complicated procedure neglects the overlapping 4s-band. The resulting Bloch wave functions and their E *vs* $\mathbf{k}$ curves are quite complicated. In the order of magnitude estimates for the

[16] G. C. Fletcher, Proc. Phys. Soc. (London) **A65**, 192 (1952).

matrix elements, $(l|\mathcal{O}|n)$, simpler expressions for the $3d$-band Bloch wave functions will be taken, namely,

$$\psi_n(\mathbf{k},\mathbf{r})=N^{-\frac{1}{2}}\sum_{\mathbf{R}} e^{i\mathbf{k}\cdot\mathbf{R}}\phi_n(\mathbf{r}-\mathbf{R}), \tag{52}$$

where $\phi_n(\mathbf{r})$ is the corresponding atomic wave function and $\mathbf{R}$ denotes the position of the lattice points. Such an approximation is not likely to introduce an order of magnitude error. Brooks[17] has shown that in such a case

$$(l|\mathcal{O}|n)\approx A(l|L_z|n),$$

where A is the ordinary one-electron spin-orbit interaction parameter for free atoms, and $(l|L_z|n)$ is the matrix element of $L_z=(1/i)(x\partial/\partial y-y\partial/\partial x)=(1/i)\partial/\partial\phi$, with respect to the angular parts of the atomic wave functions ϕ_l and ϕ_n. Most of the magnetic electrons, as is shown by Fletcher,[16] occupy the most energetic states of one of the $3d$-bands. It is assumed that all of them are in the nth band, and that its Bloch wave functions are given by Eq. (52), with $\phi_n(\mathbf{r})=yzf(r)$. This particular choice does not affect the order of magnitude, although it can alter the sign of the matrix element; thus, it is seen that a definite statement cannot be made as to the correct sign of the matrix elements, and, hence, of the final estimates of σ_1 and α_1. Bloch wave functions for the other four $3d$-bands are constructed in an analogous manner from the other $3d$-atomic wave functions, $\phi_l(\mathbf{r})$: $xyf(r)$, $zxf(r)$, $\frac{1}{2}(x^2-y^2)f(r)$, $(1/2\sqrt{3})\times(x^2+y^2-2z^2)f(r)$.

It is readily found that the only ϕ_l which has a nonvanishing matrix element of L_z with ϕ_n is the $zxf(r)$, and for this case $(l|L_z|n)=1/i$. Thus, neglecting a minus sign, since no faith can be placed in it, Eq. (50) may be written,

$$Q_{mn}=(2/\Delta E)A\langle m|\partial_y|n\rangle^*\langle m|\partial_x|l\rangle.$$

Now, in order to get a rough estimate of the order of magnitude of the matrix elements of $\partial/\partial x$, $\partial/\partial y$, they are transformed in the usual manner to matrix elements of x and y.

$$\langle m|\partial_x|n\rangle=(m/\hbar^2)(E_n-E_m)\int u_m^*xu_n d\tau_0$$

$$=(m/\hbar^2)(E_n-E_m)\langle m|x|n\rangle,$$

and similarly for $\langle m|\partial_y|n\rangle$. It should be noticed that the free electron approximation of the mth ($4s$-band) and higher bands gives zero matrix elements, since $\langle m|\partial_x|n\rangle=\langle n|\partial_x|m\rangle=0$ (u_m=constant for plane waves). The deviations of the wave functions from plane waves are thus essential for a correct calculation; the almost-free electron approximation usually gives fair estimates. However, in view of the dimensions of the unit cell and the relation $\int u_m^*u_n d\tau_0=\delta_{mn}$, it may be assumed $\langle m|x|n\rangle$ and $\langle m|y|n\rangle$ are of the order of $\sim10^{-9}$ cm.

[17] H. Brooks, Phys. Rev. **58**, 909 (1940).

Thus, Eqs. (48) and (49) may be written as

$$|\sigma_1|\sim1.0\times10^{11}\frac{A(E_n-E_m)(E_l-E_m)}{\Delta E[(E_m-E_n)^2-\hbar^2\omega^2]}M,$$

$$|\alpha_1|\sim|\sigma_1/\omega|.$$

(A factor of 10 can easily enter the last expression.) Here all bands higher than the $4s$-band are neglected, since for our frequencies they contribute much less on account of the resonant form of the denominator.

From Slater's diagram[5] the energy difference near the top of the magnetic $3d$- and $4s$-bands is estimated as

$$(E_m-E_n)\approx4\text{ ev}=6.4\times10^{-12}\text{ erg}.$$

Also, $\Delta E\approx1$ ev$=1.6\times10^{-12}$ erg, and therefore $(E_m-E_l)\approx5$ ev$=8.0\times10^{-12}$ erg. For a light wave of wavelength, say, $\lambda=6000$ A, $\hbar\omega=3.0\times10^{-12}$ erg, and

$$|\sigma_1|\sim1.0\times10^{23}AM.$$

Taking, with Brooks, the spin-orbit parameter to be $A\approx10^{-13}$ erg for nickel and $A\approx0.7\times10^{-13}$ erg for iron, and $M\approx500$ gauss for nickel, under saturation conditions at room temperature, and $M\approx1700$ gauss for iron, under the same conditions, we finally obtain:

$$\left.\begin{aligned}|\sigma_1|&\sim\ 5.2\times10^{12}\ \text{sec}^{-1}\\ |\alpha_1|&\sim\ 1.6\times10^{-3}\end{aligned}\right\}\text{for Ni,} \tag{53}$$

and

$$\left.\begin{aligned}|\sigma_1|&\sim12.4\times10^{12}\ \text{sec}^{-1}\\ |\alpha_1|&\sim\ 4.0\times10^{-3}\end{aligned}\right\}\text{for Fe.} \tag{54}$$

These values, along with the remark about the approximation of α_1, give, e.g., for the Faraday rotation in iron under saturation conditions at room temperature, an angle of $\sim280\,000°$ per centimeter.

The experimental work on these effects is not satisfactory, mainly on account of the usual difficulties that beset all optical measurements of solids. The results depend markedly on the method of preparation of the films used. All experimental work up to 1936 has been summarized by Schültz.[1] The most recent measurement of the Faraday rotation in iron is that of König,[18] who measured ϕ_F in saturated iron at room temperature for $\lambda\sim6000$ A and found it to be ($\phi_F=380\,000°$ per centimeter) approximately twice as big as previous measurements had indicated. He was able to show that the oxidation of the films can account for such differences. An analysis of the work of Foote[19] on the Kerr constants and that of Skinner and Tool[20] on the Faraday constants for $\lambda\sim6000$ A gives the following values for the quantities σ_1, α_1 within a factor of 2, taking $n=2.0$, $k=3.9$ for nickel, and $n=2.4$, $k=3.3$ for iron (see also

[18] H. König, J. Optik **3**, 101 (1948).
[19] P. Foote, Phys. Rev. **34**, 96 (1912).
[20] C. A. Skinner and A. Q. Tool, Phil. Mag. **16**, 833 (1908).

Darwin[21])

$$\left.\begin{aligned}\sigma_1 &\cong 3.1\times10^{12}\ \text{sec}^{-1}\\ \alpha_1 &\cong 1.7\times10^{-2}\end{aligned}\right\}\text{for Ni,}$$

and

$$\left.\begin{aligned}\sigma_1 &\cong 30.0\times10^{12}\ \text{sec}^{-1}\\ \alpha_1 &\cong 4.3\times10^{-2}\end{aligned}\right\}\text{for Fe.}$$

Values for different wavelengths in the visible region are not very much different. In view of the approximations necessary to get the theoretical estimates, Eqs. (53) and (54), the comparison is satisfactory.

[21] C. G. Darwin, Proc. Roy. Soc. (London) A151, 512 (1935).

5. ACKNOWLEDGMENTS

It is a pleasure to express my sincere appreciation to Professor Charles Kittel for bringing this problem to my attention and for his stimulating guidance through all my studies in the physics of the solid state. I also wish to record my appreciation of the support by the U. S. Office of Naval Research, the U. S. Signal Corps, and the International Business Machines Corporation.

PHYSICAL REVIEW VOLUME 97, NUMBER 2 JANUARY 15, 1955

Neutron Damage to the Structure of Vitreous Silica

JOSEPH S. LUKESH
General Electric Company, Knolls Atomic Power Laboratory, Schenectady, New York*
(Received September 29, 1954)

The x-ray diffraction pattern of vitreous silica has been investigated before and after exposure to neutrons. Small, but significant, changes are observed. A relation between these changes and those caused by neutron damage to crystalline forms of silica is suggested.

INTRODUCTION

SEVERAL observations have been reported on radiation damage in silica, both crystalline and vitreous. Wittels and Sherrill,[1] for instance, found that fast-neutron irradiation of a total of $\sim2\times10^{20}$ neutrons/cm^2 caused all crystalline forms to become apparently glassy with a density of 2.26. Vitreous silica also reached this density. Primak, Fuchs, and Day[2] report an increase in density and refractive index of vitreous silica. No extended study of the x-ray diffraction pattern of irradiated vitreous silica has been published. In view of the tendency of all crystalline silica materials to become glassy with a common density and of vitreous silica to increase in density to the same value, it is of interest to examine the diffraction patterns of these damaged materials in some detail. In this report, the effect of neutron irradiation on the x-ray intensity curve of vitreous silica is discussed.

EXPERIMENTAL

Vitreous silica of high purity[3] was irradiated by neutrons at the Materials Testing Reactor to an exposure of $nvt\sim2\times10^{20}$, at a temperature of about 50°C. The specimen showed no visual change other than a slight violet discoloration. The density was found to have increased from 2.21 to 2.25, in excellent agreement with previous work. Also confirming earlier observations, the refractive index (sodium D line) increased from 1.45706 ± 0.00004 to 1.46687 ± 0.00010.

The x-ray intensity curves for both unirradiated and irradiated material were obtained using a spectrometer and Geiger counter. Values were measured at intervals of one degree 2θ or less, except at high angles where larger intervals were used. Readings were taken to $\sin\theta/\lambda=0.700$, above which point all scattering is essentially incoherent. Filtered copper and molybdenum radiations were used. Since the radiation was not strictly monochromatic, no Fourier analysis has been made of the data. It is hoped that in the future, data suitable for such analysis can be obtained. The intensities were measured by using the fixed count method in which the time to register a fixed number of counts is recorded. The probable error is less than one percent.

Spectroscopic analysis of the unirradiated material is given in Table I. The results may be in error by a factor of two or three.

TABLE I. Spectroscopic analysis of vitreous silica.

Constituent	Abundance[a] (parts per million)
V	400
Cr	200
Ti	120
Mn	60
Cu	16
Mg	6
Al	6
B	$\ll$200

[a] May be in error by a factor of two or three.

* The Knolls Atomic Power Laboratory is operated by the General Electric Company under contract with the United States Atomic Energy Commission.

[1] M. Wittels and F. A. Sherrill, Phys. Rev. 93, 1117 (1953).

[2] Primak, Fuchs, and Day, Phys. Rev. 92, 1064 (1953).

[3] Provided by Corning Glass Works, Corning, New York.

RESULTS AND DISCUSSION

The results of the x-ray analysis are shown in Fig. 1. Open circles denote the unirradiated material, and the solid circles, the irradiated. Although the ordinate is arbitrary (being the recorded counts per second after correction for all factors except Compton scattering), the curves were adjusted to be equal at $\sin\theta/\lambda=0.700$. A rough quantitative relation can, thus, be assumed between the two. This relation must not be assumed too rigorously since there was a slight residual beta activity in the irradiated sample, which might tend to increase the apparent intensity measurements. However, the integrated intensities under the first peak of the two curves differ by only a few percent. The curves do not show data above $\sin\theta/\lambda=0.459$; at higher values, the intensity curves are essentially identical.

The intensity curve of unirradiated material agrees, except in minor detail, with that of Warren, Krutter, and Morningstar,[4] which was constructed from a microphotometer trace of a film made with crystal-monochromatized radiation. There is, for instance, a slight shoulder on the main peak which is not evident in their curve. However, considering the diversity of the techniques and the fact that the radiation used in the present work was not strictly monochromatic, the curves may be considered in good agreement.

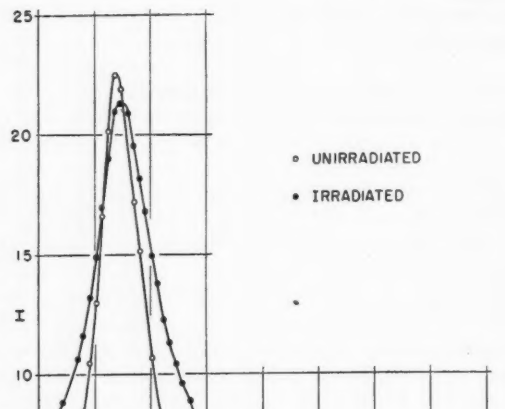


FIG. 1. X-ray intensity curves of irradiated and unirradiated vitreous silica. Vertical scale arbitrary.

The intensity curve of the irradiated vitreous silica is significantly different from that of the unirradiated. The main peak is broadened and shifted to a slightly higher value of $\sin\theta/\lambda$. The shift is from 0.120 to 0.124, and is considered to be real since every effort was made to eliminate such errors as might occur in specimen mounting, etc. The broad peak at $\sin\theta/\lambda=0.350$ to 0.400 is sharpened, and the integrated area appears to have increased. Detail in the region $\sin\theta/\lambda=0.230$ to 0.300 changed but slightly.

Although in visual appearance vitreous silica suffers little damage from neutron irradiation, there is a small but significant structural change. Wittels and Sherrill[1] suggest that all four solid forms of silica are reduced to a common vitreous phase and—more important—that there is a possibility of a solid state reaction from the vitreous phase to quartz. If such a reaction does prove to exist, particularly if it occurs from the vitreous phase regardless of the original state of the material, it would seem that the common vitreous phase must be one in which the silicon-oxygen tetrahedra have degrees of freedom not permitted in the original vitreous structure. Such freedom could be obtained by breaking of Si-O bonds, a condition that should be made evident by a decrease in the area under the first peak of the radial distribution curve of vitreous silica. If this is the case, neutron bombardment may be considered as playing a role similar to that of chemical fluxes in breaking down the tightly bonded silicon-oxygen framework.

ACKNOWLEDGMENTS

Density measurements were made by H. G. Sowman, refractive indices by W. O. Haas, and spectroscopic analysis by F. P. Landis of this laboratory.

[4] Warren, Krutter, and Morningstar, J. Am. Ceram. Soc. 19, 202 (1936).

Photostimulated Emission of Some Activated Alkali Halide Phosphors*

C. E. MANDEVILLE AND H. O. ALBRECHT
Bartol Research Foundation of The Franklin Institute, Swarthmore, Pennsylvania
(Received August 30, 1954)

The intensity of the photostimulated ultraviolet emission of x-ray excited phosphors of NaCl-Tl, KCl-Tl, NaCl-Ag, and KCl-Ag has been measured as a function of the wavelength of the incident light. The results indicate that, in the first thirty days following excitation at least, *F* centers are primarily responsible for the storage of energy. Thus, the storage properties depend upon the nature of the host material, rather than upon the intentionally added activator. Evidence is also presented which seems to indicate the presence of a deep trap associated with the silver ions of KCl-Ag which has a period of decay comparable to that of the *F* centers. Excitation, storage, and all measurements were carried out at 22°C. Polycrystalline materials, prepared by the writers, were used throughout the investigation.

INTRODUCTION

THE photostimulated emission of activated phosphors gives a measure of the "light storage" which is in turn proportional to the number of electrons stored in traps of the solid. In order to ascertain the amount of storage, a procedure is followed in which excited phosphors are irradiated by light, and the resulting photostimulated emission of "stored light" observed. Temperature is a parameter in these measurements, because both electron trapping and the efficiency of the luminescence centers for conversion of potential energy into light may vary as the temperature is changed. Consequently, to obtain a true measure of storage, excitation, storage of the excited materials in darkness, and subsequent photostimulation —all operations—were carried out at room temperature (22°C). In the studies to be described in this paper, the photostimulating light had a wavelength, λ, of more than 3000A, and the stimulated emission a wavelength of $\lambda' \sim 2500$A. Photosensitive G-M tubes were employed as detectors, having a maximum of spectral response at 2500A and no detectable response for values of λ of 3000A or more. Since the detectors did not respond to the photostimulating light, it was possible to observe the stimulated emission while the photostimulating light was acting. Investigations employing this technique have been previously described by the writers.[1,2] The measurements of the present paper differ from the earlier observation[1,2] in that the intensity of the photostimulated emission has been measured as a function of the wavelength of the photostimulating light. From the earlier measurements it was evident that since the stored energy had a very long period of storage, relatively large trap depths must be involved. Furst and Kallmann[3] have studied in particular NaCl-Ag and have given some qualitative estimates of the depths of the traps concerned with storage. They have also speculated[4] as to whether the traps responsible for the long period storage might be negative ion vacancies in the crystal lattice, forming *F* centers when trapping electrons. They also noticed that the extent of coloration of crystals of NaCl-Ag appeared to be directly related to their response to photostimulation.

The various measurements of energy storage which have been carried out by Kallmann *et al.*[3,4] have consisted for the most part in observation of a post-stimulation phosphorescence; that is, a phosphorescent afterglow was observed after the photostimulating light was extinguished. Their detector was a photomultiplier tube (RCA-1P28) which responded not

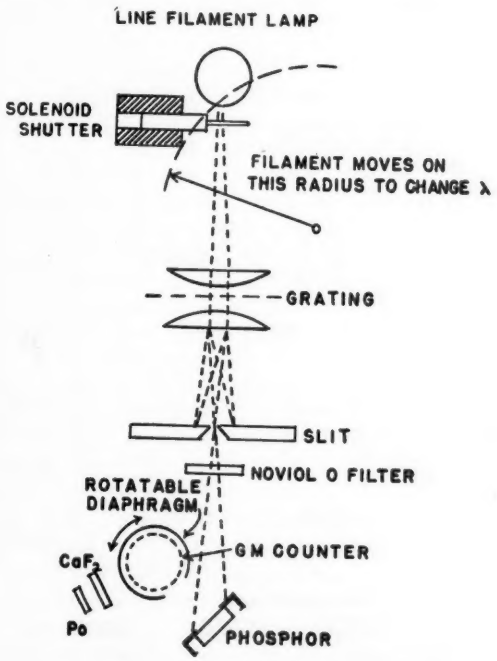


FIG. 1. Schematic diagram of the apparatus.

* Assisted by the joint program of the U. S. Office of Naval Research and the U. S. Atomic Energy Commission.
[1] C. E. Mandeville and H. O. Albrecht, Phys. Rev. **91**, 566 (1953).
[2] H. O. Albrecht and C. E. Mandeville, J. Franklin Inst. **257**, 353 (1954).
[3] M. Furst and H. Kallmann, Phys. Rev. **91**, 1356 (1953).
[4] H. Kallmann and M. Furst, Phys. Rev. **83**, 674 (1951).

only to the photostimulated emission but to the stimulating light as well; therefore, it was possible to observe only the phosphorescent emission which persisted after the photostimulating light was turned off.

In the present investigations, light from a tungsten lamp was passed through a Coleman monochromator and allowed to fall upon an excited phosphor adjacent to which was placed a photosensitive G-M tube in which was recorded, as a function of the wave length of the incident light, the photostimulated ultraviolet emission. The monochromator supplied photons ranging in wavelengths from 3000A to 7000A in a channel of constant width 100A. The number of photons lying in the channel was determined at each setting of the wavelength by a thermopile calibration. A schematic diagram of the apparatus used for the photostimulation studies is shown in Fig. 1. The constancy of the sensitivity of the counter was frequently checked by counting ultraviolet from Po-alphas on fluorite.

The purpose of the investigation was to determine the nature of the traps responsible for the relatively long-period energy storage. Consequently, measurements were commenced only after sufficient time had elapsed that any phosphorescence arising from thermal agitation at room temperature was no longer detectable. Thus, the relatively shallow traps had emptied so that the results cannot be compared with other available data.[5,6]

THE MEASUREMENTS

The photostimulated ultraviolet emission of two phosphors is plotted in Fig. 2 as a function of the wavelength of the incident light. Figure 2(A) applies to NaCl-Tl (thallium chloride concentration 0.1 mole percent) and Fig. 2B to KCl-Tl of the same mole percent concentration of activator. Both samples had received the same x-ray dosage, 65r. This dose was applied in a time of 10 seconds and the excited phosphors were immediately stored in darkness. At various times after excitation, as indicated on the curves, the photostimulated emission was observed. The practice was usually followed of commencing in the red and proceeding in steps of 100A to 250A through the peak of the curve to shorter wavelengths.

In the case of both curves of Fig. 2, the maximum of photostimulated emission occurs at the peak of the *F* band of the host material. Many investigations of *F* bands by earlier workers using the more conventional absorption methods have located the maximum of absorption in the *F* band at 4750A for NaCl and 5750A for KCl. The half-widths of the observed curves of Fig. 2 are in agreement with those of the *F* bands found by the absorption method. Thus, the curves of Fig. 2 may be interpreted as showing that *F* centers are destroyed by the photostimulating light, the released electrons entering the conduction band whence they proceed to the filled band by way of a luminescence center; that is, recombination with a positive hole occurs at the luminescence center where the ultraviolet light is also emitted. The ultraviolet emission

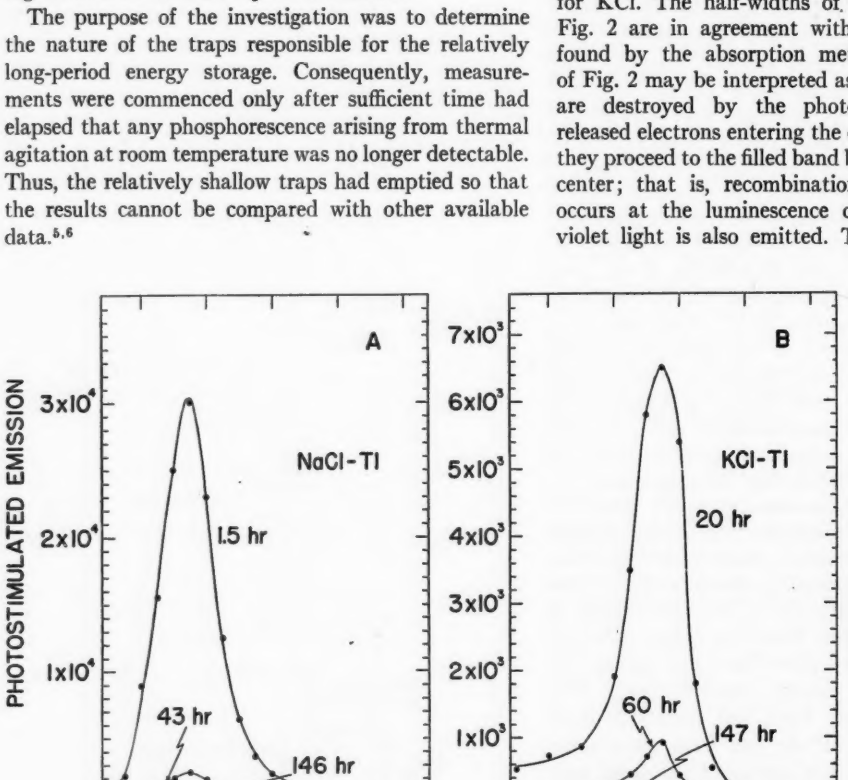


FIG. 2. Photostimulated emission of thallium-activated NaCl and KCl (TlCl concentration 0.1 mole percent) as a function of wavelength. The location in λ and widths of the peaks are identical with the *F* band observed in conventional absorption studies. Excitation, storage, of samples in darkness, and all measurements at 22°C.

[5] W. Buenger and W. Flechsig, Z. Physik 69, 637 (1931).
[6] P. D. Johnson and F. E. Williams, J. Chem. Phys. 21, 125 (1953).

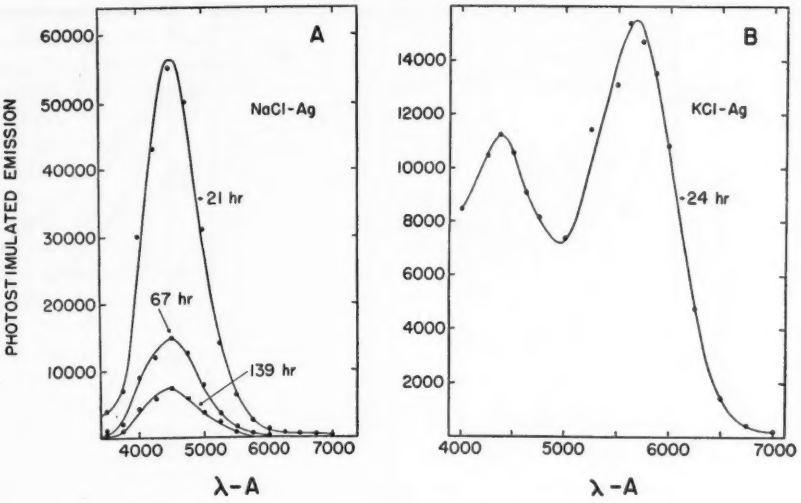


FIG. 3. Photostimulated emission of silver-activated NaCl and KCl (AgCl concentration 0.1 mole percent) as a function of wavelength. In addition to peaks corresponding to the *F* band of either host crystal, a strong maximum at 4500A is detected in KCl-Ag. Excitation, storage, and all measurements at 22°C.

was observed at each point of the curves of Fig. 2 simultaneously with application of the photostimulating light. The time of photostimulation at each point was usually a period of a few seconds. The time of completing an entire curve was about one hour.

It is clear from the curves of Fig. 2 that the long-period storage of energy in the thallium-activated alkali halides arises from the trapping of electrons to form *F* centers.[7]

The photostimulated emission of NaCl-Ag and KCl-Ag (AgCl concentration 0.1 mole percent) is shown in Figs. 3(A) and 3(B). In this case, NaCl-Ag was x-rayed for five seconds (33r), and KCl-Ag for five minutes (2000r). The time after initial excitation at which each curve was obtained is specified in the figure. Here again the *F* bands of the host crystals are clearly in evidence.[8] However, in the case of KCl-Ag, an additional band peaked at 4500A is present with a considerable intensity. As indicated by Fig. 2(B), this intense band is not present in thallium activated KCl. It is reasonable to assume, therefore, that this short-wave band is in some way related to the presence of the silver ion.

To obtain information as to any possible relationship between the two peaks of intensity of photostimulated emission of KCl-Ag, the excited phosphor of Fig. 3(B) was bleached at 6500A, and the relative intensities of the two bands were determined at their respective peaks as the bleaching proceeded. The total number of photostimulated counts recorded during bleaching was taken as a measure of the extent of bleaching. Short bursts of photostimulated ultraviolet emission were drawn from the phosphor first at the peak at 5750A, then at the peak at 4500 A, to measure the heights of those peaks as bleaching at 6500A continued. The bleaching was halted, of course, for a time sufficiently long to permit these intensity measurements. During a bleaching time of 62 minutes, a total of 1.28×10^6 counts of photostimulated ultraviolet emission was recorded in the photosensitive G-M tube. The rate of decrease of the peak value of the *F* band and of the band centered at 4500A is shown in Fig. 4(A), the bleaching being carried out at 6500A. The intensity of the short wave peak was clearly relatively unaffected, but the *F* band peak diminished rapidly. A similar curve is given in Fig. 4(B) where the respective heights of the two peaks are plotted as exhaustion proceeds by photostimulation at 4000A. In this case, both peaks were bleached as might be expected, since the photostimulating light was sufficiently energetic to reduce either peak. Since, as indicated by the curve of Fig. 4(A), bleaching in the *F* band does not reduce the intensity of the peak at 4500A, it must be concluded that the photostimulated emission at 4500A must correspond to some type of trap other than *F* centers; that is, the 4500A peak in KCl-Ag appears to be unrelated to the *K* band of pure KCl.[7] Further confirmation of this point lies in the obvious fact that no peak at 4500A of an intensity comparable to that of KCl-Ag is found in KCl-Tl [Fig. 2(B)].

To examine further the nature of the maximum of of photostimulated emission at 4500A in KCl-Ag, additional samples were prepared having AgCl concentrations of 0.006 mole percent and 0.2 mole percent respectively. Curves of the photostimulated emission from these materials are shown in Fig. 5(A) and Fig.

[7] Examination of the curves to the blue of the *F* band in each case shows that there is a greater amount of photostimulated emission resulting from irradiation of KCl-Tl in that spectral region than from NaCl-Tl. This increased emission may be related to the *K* band reported in absorption studies of x-rayed KCl [see F. G. Kleinschrod, Ann. Physik **27**, 86 (1936)].

[8] The curves of Figs. 2(A), 2(B), and 3(A) show a reduction of peak intensity with time which is partly fictitious. The photostimulating light was sufficiently intense that the measurements themselves appreciably decreased the number of stored electrons.

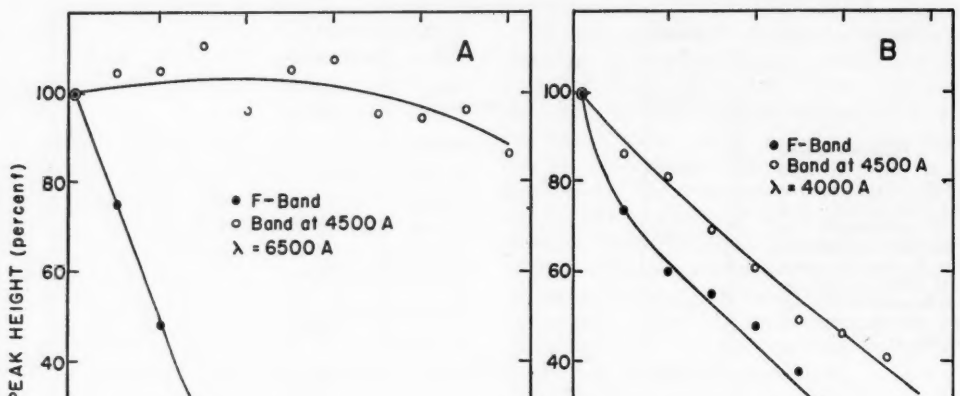


FIG. 4. Heights of the peaks of photostimulated emission of KCl-Ag as a function of bleaching time; curves *A*, at 6500A, curves *B* at 4000A. All measurements at 22°C.

5(B) where it is clear that the peak of emission at 4500A decreases relative to the *F* band peak with decreasing silver concentration, further substantiating the view that the short-wave peak of KCl-Ag derives from the presence of Ag ions in the crystal lattice. A possible explanation is that quanta of wavelength 4500A are sufficiently energetic to remove trapped electrons from a long-lived metastable state to a higher state from which a return of the electron to the light-emitting levels of the luminescence center is allowed.

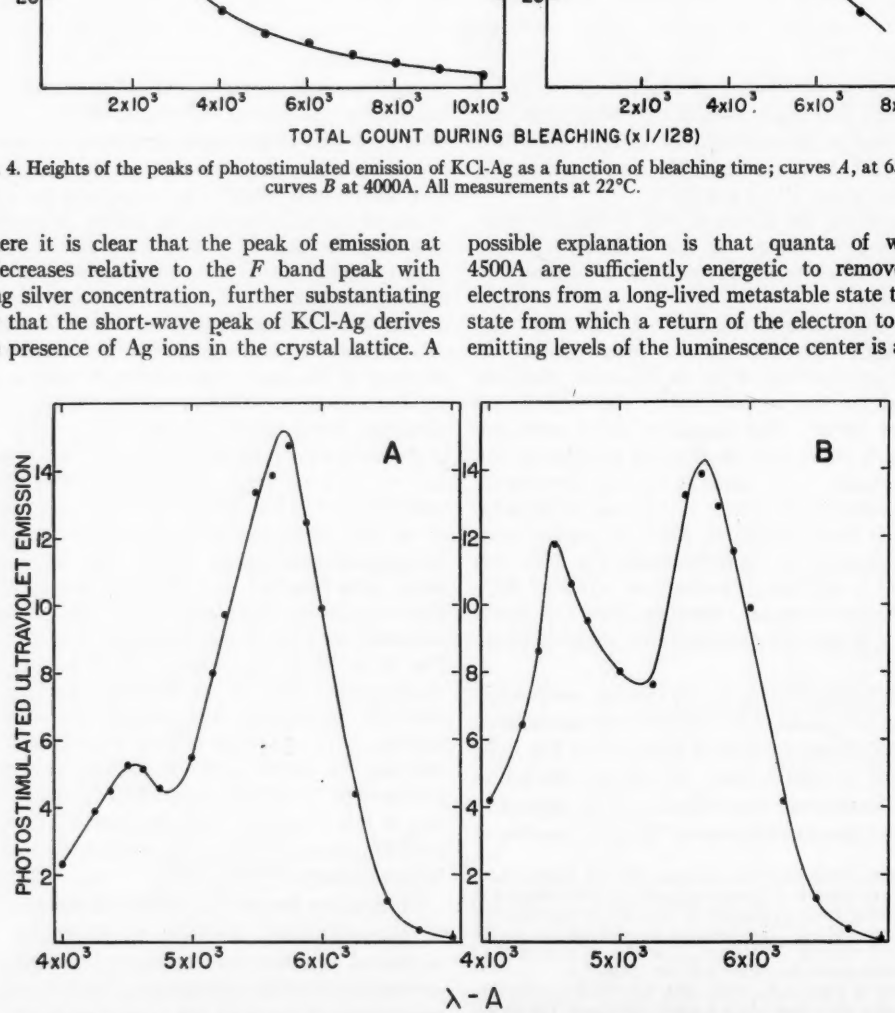


FIG. 5. Photostimulated emission of KCl-Ag. Curve *A*, mole concentration of AgCl 0.006 percent; Curve *B*, 0.20 percent. These curves were obtained 24 hours after initial excitation.

In Fig. 6 are plotted decay curves of the stored energy in KCl-Ag, KCl-Tl, NaCl-Ag, and NaCl-Tl. It is evident from the curves that the decay is essentially the same for a given host crystal, irrespective of the nature of the activator ion. Thus, the differences in storage times relate purely to the variations in host material. In Fig. 6, all samples of KCl, thallium-activated or silver-activated, decay with approximately the same slope as do those of NaCl. However, it is also clear that storage is more long-lasting in NaCl than in KCl. Not shown in the figure, a single point taken for each curve at a time of forty days left unchanged the conclusions above, which are based upon behavior during the first twenty days as plotted in Fig. 6. To obtain the decay curves of Fig. 6, a short burst of photostimulated emission of duration six seconds was drawn from the phosphor at the indicated intervals of time. The intensity of the photostimulated emission was taken to be proportional to the number of electrons remaining stored in traps in the solid. It has often been suggested that long-period storage may depend upon the nature of the activator ion. For example, in a previous publication,[2] some results of the writers seemed to indicate that the stored energy in KCl-Tl decayed more rapidly than that in KCl-Ag and similarly for thallium- and silver-activated NaCl. The results were in error, because too intense a photostimulating light was employed; that is, the observation of the number of stored electrons destroyed too many F centers, giving the impression of a rapid decay of the stored energy in the thallium-activated materials. In the present measurements, a very faint photostimulating light was employed and tests were performed which showed a negligible loss to arise from the measurements themselves. In observing the decay of energy storage in KCl-Ag, the reduction with time of both peaks was followed as indicated. The supposed metastable state appears to decay with about the same period as do the F centers. There is always the possibility that the metastable level is "fed" by electrons which have been removed from the F centers.

The luminescent materials were produced at Bartol by fusion in a platinum crucible and subsequent rapid cooling on an aluminum plate. A polycrystalline structure resulted. No single crystals were employed.

CONCLUSION

From the data presented in the foregoing sections, it is clear that the long-period storage of energy in the activated alkali halides results from the trapping of electrons to form F centers in the host material. The

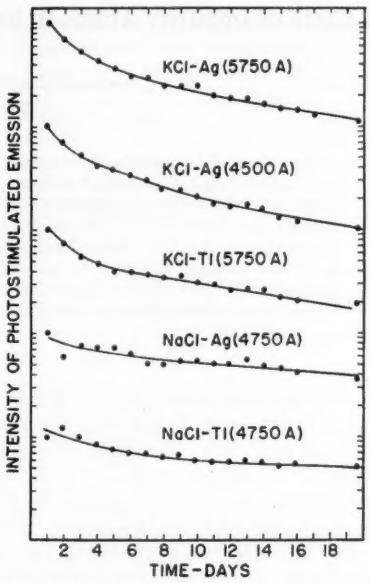


FIG. 6. Decay of stored energy in thallium- and silver-activated alkali halides. These curves show that the rate of decay depends primarily upon the nature of the host crystal, not upon the type of activator. Excitation, storage, and all measurements at 22°C.

decay curves of Fig. 6 can be related to the expression[9]

$$n = n_0 \exp(-sle^{-E/kT}),$$

where n is the number of trapped electrons, s a constant characteristic of the host crystal, and E the thermal dissociation energy of the F centers. The mean life of the F centers of NaCl has been calculated to be much greater than 86 days,[9] whereas the curves of Fig. 6 indicate a much shorter lifetime.[9]

It is to be remembered that the measurements under discussion pertain to those traps which are dominant in storage over the first thirty days following initial excitation.[10] The possibility of existence of still deeper traps of lesser population remains. Study of any such traps must await an extended aging of the x-rayed materials.

Note added in proof.—The peak of photostimulated emission of KCl-Ag centered at 4500A may correspond to the E band of recent absorption measurements. See H. W. Etzel and J. H. Schulman, J. Chem. Phys. **22**, 1549 (1954).

[9] F. Seitz, Revs. Modern Phys. **26**, 7 (1954). See Sec. 29, "Bleaching," which also includes a discussion of reasons for the observed unduly rapid early decay of F centers.

[10] For remarks concerning the effects of retrapping upon the observed lifetime of F centers, see R. S. Alger, J. Appl. Phys. **21** 30 (1950).

Ground State of Impurity Atoms in Semiconductors Having Anisotropic Energy Surfaces*

MURRAY A. LAMPERT
RCA Laboratories, David Sarnoff Research Center, Princeton, New Jersey
(Received September 22, 1954)

The nature of the ground state of a substitutional impurity atom in a crystal having an anisotropic energy-band structure is re-examined. A variational calculation has been made for the case of a band whose energy contours consist of several, symmetrically-located ellipsoids. Unlike the previous scalar, hydrogenic impurity model, this calculation at the outset uses the experimentally determined effective-mass-tensor components. Results for the impurity binding energy in germanium and silicon are on the order of thirty percent lower than those obtained directly from activation-slope measurements. Further, theoretical agreement with the experimentally determined critical impurity density for vanishing binding energy in germanium is much improved over the older model, though still not perfect.

IT has been customary to describe the ground state of a Group III or Group V impurity atom, substitutionally introduced into the diamond-type lattice (germanium or silicon) in terms of a hydrogen-like model[1] in which the electron has a scalar effective-mass. However the recent cyclotron resonance experiments[2] have revealed that electrons at the bottom of the lowest conduction band in both germanium and silicon have highly anisotropic effective-mass tensors. Further, for both crystals the E *vs* $\mathbf{k}$ map for the lowest conduction band exhibits several identical, symmetrically located energy valleys, the number and their axial locations being indicated in the first two columns of Table I. The measured effective-mass values corresponding to principal axes of the constant-energy ellipsoids are given in the third and fourth columns. It is evident that a reconsideration of the theoretical problem is required.

A formalism sufficiently powerful, in principle, to handle the impurity problem in the presence of the above complications has recently been published by Koster and Slater.[3] In this formalism a set of simultaneous difference equations is obtained relating Fourier-expansion coefficients of the Bloch-wave energy $E(\mathbf{k})$ to Wannier-function-expansion coefficients of the impurity wave function. In the case of a single minimum $\mathbf{k}_0$ in the E *vs* $\mathbf{k}$ map, passage from the difference equations to the standard effective-mass Schrödinger equation is achieved via a Taylor expansion of $E(\mathbf{k})$ about $\mathbf{k}_0$ and some other simplifying assumptions. Where the E *vs* $\mathbf{k}$ map exhibits several equal minima, in a first approximation one may follow the same procedure and Taylor-expand $E(\mathbf{k})$ about one minimum, thereby obtaining the effective-mass Schrödinger equation and its solution for that particular energy valley. The correct impurity wave function is then a superposition[4] of the individual wave functions for each energy valley. Because the latter functions are "building blocks" for the correct wave function it was felt worthwhile to study them and also thereby obtain a first approximation to the binding energy of the ground state using the experimental mass values.

The effective-mass Schrödinger equation for one energy valley, with the reduced Coulomb potential $-e^2/\epsilon r$, may be written, after a change to dimensionless units, as

$$H\psi=-\left\{\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{1}{C}\frac{\partial^2}{\partial z^2}+\frac{1}{r}\right\}\psi=\xi\psi, \qquad (1)$$

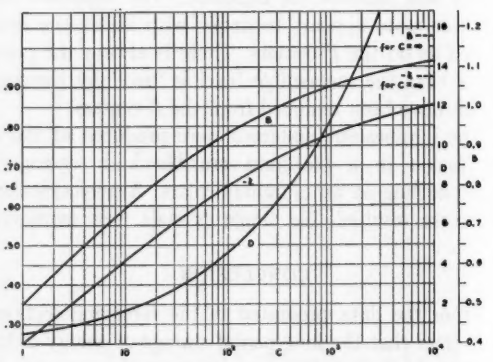


FIG. 1. "Pancake" wave-function parameters *vs* $C=m_l/m_t(C\geq 1)$.

* This work was supported by the Signal Corps of the U. S. Army.

[1] Wartime contributions to this problem are reviewed in H. C. Torrey and C. A. Witmer's *Crystal Rectifiers* (McGraw-Hill Book Company, Inc., New York, 1948). More recent work is reviewed by G. W. Castellan and F. Seitz in *Semiconducting Materials* (Proceedings of the Reading Conference) (Academic Press, New York, 1951). In the hydrogen-like model for the impurity, the ground state has the radius $a_I=(\epsilon m/m^*)a_0$ and the binding energy $E_I=(m^*/\epsilon^2 m)E_{IH}$. Here ϵ is the dielectric constant of the host crystal, m the free electron mass, m^* the effective mass of the lowest-conduction-band electrons for Group V impurities (or of topmost-valence-band holes for Group III impurities) a_0 is the Bohr radius, $a_0=0.53$ A, and $E_{IH}=e^2/2a_0$, e being the electronic charge.

[2] Lax, Zeiger, Dexter, and Rosenblum, Phys. Rev. **93**, 1418 (1954). Also B. Lax, paper read at the Amsterdam Conference on Semiconductors, June 1954, Physica (to be published).

[3] G. F. Koster and J. C. Slater, Phys. Rev. **95**, 1167 (1954), particularly Eqs. (8), (3), (6), and (12).

[4] In superposing these individual wave functions account must be taken of their nonorthogonality—they are coupled together by the Coulomb potential. A study of this coupling has been made by J. Luttinger and W. Kohn. For clarification of this aspect of the problem I am indebted to J. Luttinger (private communication).

TABLE I. Experimental and theoretical results for conduction-band electrons and donor impurities in germanium and silicon.

	No. of Min.	Axial location	m_l/m^a	m_t/m^a	E(exp)[b] ev	E(calc) ev	R_t A	R_l A	m^*/m	a_I A
Ge	8(or 4)[c]	⟨1,1,1⟩	1.4	0.083	As:0.0127	0.0089	64.6	23.9	0.24	35
Si	6	⟨1,0,0⟩	0.99	0.19	P: 0.043	0.0285	24.7	13.5	0.46	14

[a] Cyclotron-resonance data reported by B. Lax at the Amsterdam Conference on Semiconductors, June, 1954, Physica (to be published).
[b] Activation-energy data reported by J. A. Burton at the Amsterdam Conference on Semiconductors, June, 1954, Physica (to be published).
[c] Whether there are $2M$ or M energy valleys depends on whether the minima are located in the interior of the first Brillouin zone or on boundary faces of the zone.

with $C=m_l/m_t$ and $r=(x^2+y^2+z^2)^{\frac{1}{2}}$. The coordinate axes are chosen to coincide with principal axes of the constant-energy ellipsoids for the chosen energy valley. The units of length and energy are, respectively, $a_t=a_I/2$ and $E_t=4E_I$,[1] with $m^*=m_t$.

For $C\neq 1$, (1) is not separable and in consequence a variational method has been employed to estimate the binding energy ξ of the ground state. The exact hydrogenic ground-state wave function, $\psi_{(C=1)}=(8\pi)^{-\frac{1}{2}}\times\exp[-(\frac{1}{4}x^2+\frac{1}{4}y^2+\frac{1}{4}z^2)^{\frac{1}{2}}]$, suggests a simple trial wave function for the general case, $C\neq 1$, which one would not only expect to represent the true solution quite well, but which also proves quite tractable in calculations, namely the normalized function

$$\psi=\left(\frac{B^2D}{\pi}\right)^{\frac{1}{2}}\exp[-(B^2x^2+B^2y^2+D^2z^2)^{\frac{1}{2}}]. \quad (2)$$

Minimization of the energy integral $I=\int\psi^*H\psi d\tau$ [where H is given by (1)] with respect to B and D results in the following set of equations, valid for the case $C\geq 1$, involving an auxiliary variable u:

$$C=\tfrac{1}{2}(u^2+1)\left[\frac{u^3}{u-\tan^{-1}u}-(u^2+1)\right], \quad (3)$$

$$D=\tfrac{3}{2}C\frac{u-\tan^{-1}u}{u^3}, \quad B=\frac{D}{(u^2+1)^{\frac{1}{2}}}, \quad (4)\text{–}(5)$$

$$-\xi\simeq -I=\frac{D}{u}\tan^{-1}u-\tfrac{3}{4}B^2-\frac{1}{3}\frac{D^2}{C}. \quad (6)$$

By first plotting C, D, B, and $-\xi$ against u, one then obtains the desired plots of D, B, and $-\xi$ as functions of C, shown for the range $1\leq C\leq 10^4$ in Fig. 1. For the case $C<1$, which is not at present of experimental interest, a set of equations somewhat like the set (3)–(6) is obtained, with $\tan^{-1}$ replaced by $\tanh^{-1}$.

With $C>1$, the wavefunction (2) is a "pancake" function. Plausible measures of the transverse and longitudinal spreads of this "pancake" function are, respectively, $R_t=a_t/B$ and $R_l=a_t/D$.

The experimental values for the binding energy E and the calculated values of $E=(-\xi)E_t$, R_t, and R_l are presented, for germanium and silicon, in Table I. Also included are the scalar effective-mass ratios m^*/m of the "old" theory, as determined from E(exp), and the corresponding ground-state radius a_I.

For electrons in the lowest conduction band in silicon ($\epsilon=12$), $C=5.2$, $a_t=16.7$ A, $E_t=0.0713$ ev and from Fig. 1, $D=1.24$, $B=0.675$, and $-\xi=0.40$. E(calc) is 34 percent lower than E(exp) corresponding to phosphorus impurity.

For electrons in the lowest conduction band in germanium ($\epsilon=16$), $C=16.9$, $a_t=51$ A, $E_t=0.0175$ ev and from Fig. 1, $D=2.13$, $B=0.79$, and $-\xi=0.508$. E(calc) is 30 percent lower than E(exp) corresponding to arsenic impurity.

As pointed out above, the impurity ground state is a superposition of the variationally-determined "pancake" functions, each with the spatial orientation determined by the constant-energy ellipsoids for the particular energy valley. The resultant ground state is a highly scalloped wave-function whose spatial extent, for purposes of discussion of the "overlapping impurities" problem may be taken as R_t. "Cubic" packing of impurities would correspond to a density of $(2R_t)^{-3}$ cm^3, or 4.6×10^{17} cm^{-3} for germanium (Table I). This is much closer to the published data of Debye and Conwell[5] which indicates a vanishing impurity ionization-energy at a donor concentration of 1.6×10^{17} cm^{-3}, than the result predicted by the "old" scalar model. The "cubic" packing density in the latter model is $(2a_I)^{-3}$ cm^{-3}, or 2.8×10^{18} cm^{-3} (Table I).

[5] P. P. Debye and E. M. Conwell, Phys. Rev. 93, 705 (1954), Fig. 12.

Effect of Crystal Distortion upon Change of Resistivity of Silicon by Heat Treatment

W. C. DASH
General Electric Research Laboratory, Schenectady, New York
(Received September 27, 1954)

Silicon crystals grown by the Teal-Little method under certain conditions change resistivity upon heat treatment at about 450°C, becoming more *n*-type. When crystals of this type are bent at 1100°C and quenched, heat treatment at 450°C indicates that the distorted regions are more stable than undistorted regions. This suggests that structural defects retard the appearance of *n*-type carriers and may be responsible for variations among samples grown under different conditions.

WHEN silicon crystals grown by the Teal-Little method are heated for a period of from 1–50 hours at a temperature of the order of 450°C, large changes in resistivity are frequently produced.[1] In effect, a number of *n*-type carriers of the order of 10^{16} per cubic centimeter are introduced throughout most of the bulk of the crystal by prolonged heat treatment. The phenomenon is found primarily in crystals which have been rotated during growth, unrotated ingots being much more stable. The difference between the two is presumed to be caused by variations in the growth rate of the rotated crystal resulting from azimuthal inhomogeneities in temperature of the melt. Apparently diffusion of surface impurities is not involved in the resistivity change. For silicon which changes resistivity in this way (crystals rotated at 60 rpm during growth were used here), evidence has been obtained that plastic distortion at high temperature changes its heat treatability at 450°C in such a way that the most distorted regions are most stable.

In order to study the effect of plastic distortion on heat treatability, specimen bars of single-crystal silicon $\frac{1}{8}$ inch×$\frac{1}{8}$ inch×1 inch were bent[2] at 1100°C in a nitrogen atmosphere. The specimens were placed between two matching quartz blocks, one concave and one convex with the desired radius. The bending was done by forcing the blocks together slowly.

The specimens were then removed quickly from the furnace and set onto a quartz plate at room temperature. Rapid cooling was found to be essential to prevent inhomogeneities in resistivity after bending.

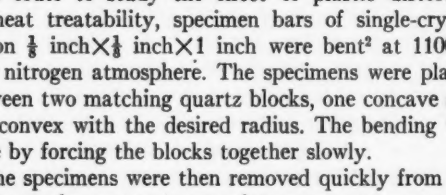


FIG. 1. Simplified sketch of a bent silicon sample which has been heat-treated at 450°C for a time sufficient to show *p-n* barriers between areas of low and high distortion. In an actual case slip bands and other inhomogeneities cause a jagged boundary between the *p*- and *n*-regions.

[1] Fuller, Ditzenberger, Hannay, and Buchler, Phys. Rev. 96, 833 (1954).
[2] C. J. Gallagher, Phys. Rev. 88, 721 (1952).

The ends of a bent bar and a "neutral plane" region in the center were relatively undistorted. This was verified by etching and observing the pattern of etch pits formed. The resistivity of a bar cut from the most distorted region was found to be the same as that of a bar from the neutral plane region within the accuracy of measurement, provided that the bar was cooled rapidly from bending temperature and received no further heat treatment.

A typical *p*-type specimen of about 15 ohm-cm initial resistivity was bent to a radius of about two centimeters and then heated for one hour at 450°C in air. Titanation or pulse-plating with copper showed a conversion of the neutral plane region and the ends of the bar to *n*-type in a pattern approximating the pattern of etch pits seen before heat treatment. Figure 1 shows a simplified sketch of the conductivity type for different parts of the specimen after heat treatment at 450°C. Further heating for $1\frac{1}{2}$ hours at 450°C converted the entire bar to *n*-type. Polishing and re-etching showed an apparently unchanged etch-pit density and distribution. Studies made on control samples at 450°C indicated that copper diffusion, found important in germanium,[3] was not important in these experiments.

Similar qualitative results were obtained for specimens from several different ingots bent to a number of different radii ranging from 1 cm to 1 meter. It was found in general that the smaller the radius of the bend (i.e., the greater the distortion) the longer the time required to produce complete conversion from *p*- to *n*-type. At present no quantitative relationship is known for radius of bend *versus* time of heat treatment necessary to obtain a given change in carrier concentration. Such information is very difficult to obtain because of irregularities in the deformation.

These experiments give evidence that the presence of structural defects, perhaps vacancies or dislocations, retard the appearance of *n*-type current carriers when a sample is heat treated at 450°C. This suggests that differences in heat treatability between rotated and unrotated ingots made from the same lot of silicon might result from differences in the number of structural defects.

[3] C. S. Fuller and J. D. Struthers, Phys. Rev. 87, 526 (1952).

Temperature Dependence of Electron Mobility in AgCl†

FREDERICK C. BROWN
Reed College, Portland, Oregon
(Received August 13, 1954)

By using improved crystal counter techniques, electron mobility has been investigated as a function of temperature in AgCl. The measurements were made on annealed samples carefully grown from the melt in which the range of conduction electrons was of the order of 10^{-4} cm per volt/cm of electric field strength. At high fields electron trapping in the volume of the crystal is less important and observed mobility is shown to become a constant independent of field. At 86°K, drift mobility was found to be 274 cm²/volt sec and was reproducible to within 10 percent for several samples. The data on mobility can be fitted down to 86°K by an expression of the form $\mu=2.54\times10^{5}T^{-\frac{1}{2}}$, which would suggest interaction of the electrons mainly with acoustic vibrations of the lattice. However, there exists the possibility of scattering by impurities or other imperfections which in combination with optical scattering might also lead to the observed results. Strains play an important role in this material and are shown to be associated both with shallow 0.1-ev traps and with deeper traps. The average energy for production of one electron-hole pair by beta rays is 7.5±0.5 ev. The range of holes in the samples tested, from 86°K to 150°K, was at least less than $\frac{1}{10}$ that of electrons.

I. INTRODUCTION

SEVERAL important advances have recently been made in the theory of the behavior of slow electrons in polar substances.[1–5] It has been emphasized[3] that the results of this theory are applicable only for temperatures considerably below the Debye temperature corresponding to longitudinal optical vibrations, which is from 200 to 370°K for most ionic crystals. Except for the recent work of Redfield on Hall mobility in the alkali halides,[6] there have been few experiments carried out at low temperature to compare with theory. Haynes and Shockley[7] have made measurements near room temperature of Hall and drift mobility in single crystal AgCl. They have shown that in properly prepared samples the mean free time before trapping of photoelectrons is several microseconds, giving rise to long drift paths in an applied electric field. Similar long electron ranges at low temperature make it possible to detect single ionizing particles with the AgCl crystal counter,[8] whereas in NaCl the ranges are quite short.[9] AgCl, with a Debye temperature of about 280°K, is a convenient material upon which to make measurements of mobility at low temperature. In addition, interest centers on the silver halides in connection with the photographic process and properties of the solid state.[10]

The early crystal counter work with AgCl did not yield characteristic or even very reproducible values of electron mobility.[11] However, by measuring rise times, Yamakawa[12] arrived at comparable values of mobility in two samples of AgBr at 77°K. Using improved crystal counter techniques, we will show that it is possible to get consistent results on AgCl provided that trapping and the effect of strains in the crystal can be minimized. Preliminary results have been given in a letter comparing low-temperature Hall and drift mobility in AgCl.[13]

II. THEORY OF THE METHOD

Principles of operation of the conductivity crystal counter have been treated by several authors.[14] We are interested here primarily in the shape of the conductivity pulses. The annealed crystal of AgCl with shorted electrodes is slowly cooled to at least −75°C to minimize ionic conductivity. As shown in Fig. 1, a collecting voltage V_c can then be applied to the parallel electrodes (a) and (b) through an isolating resistance R. The capacitance, C, represents the combined capacity of crystal, leads, and preamplifier input. Ionization in the form of electron-hole pairs is produced near (a) by single beta particles which impinge upon the counter from below and stop in the crystal. With the top electrode (b) positive, the secondary or conduction electrons are drawn into the crystal, inducing a small charge, q, on the capacitance, C. Under conditions of large input

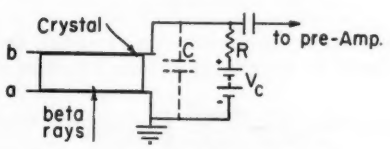


FIG. 1. Basic crystal counter circuit.

† Partially supported by grants from the National Science Foundation and the Research Corporation.

[1] H. Fröhlich and N. F. Mott, Proc. Roy. Soc. (London) **A62**, 446 (1949).
[2] Fröhlich, Pelzer, and Zienau, Phil. Mag. **41**, 221 (1950).
[3] F. Low and D. Pines, Phys. Rev. **91**, 193 (1953); also Phys. Rev. **90**, 297 (1953).
[4] E. P. Gross, Technical Report 55, Laboratory for Insulation Research, Massachusetts Institute of Technology, 1952 (unpublished).
[5] F. Seitz, Revs. Modern Phys. **26**, 7 (1954), see pp. 23–24.
[6] A. G. Redfield, Phys. Rev. **94**, 537 (1954).
[7] J. R. Haynes and W. Shockley, Phys. Rev. **82**, 935 (1951).
[8] P. J. van Heerden, thesis Utrecht, 1945 [see Physica **16**, 505 (1950)].
[9] H. Witt, Z. Physik **128**, 442 (1950).
[10] F. Seitz, Revs. Modern Phys. **23**, 328 (1951).
[11] R. Hofstadter, Nucleonics **4**, 4, 2 (1949); **4**, 5, 29 (1949).
[12] K. A. Yamakawa, Phys. Rev. **82**, 522 (1951).
[13] F. C. Brown, Phys. Rev. **92**, 502 (1953); see also Phys. Rev. **92**, 858(A) (1953).
[14] A. G. Chynoweth, Am. J. Phys. **20**, 218 (1952). See also references 8 and 11.

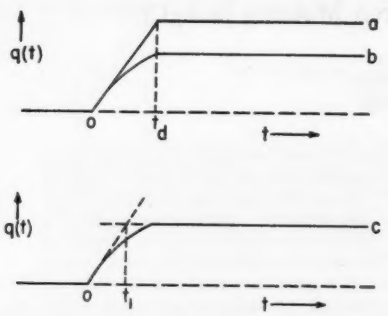


FIG. 2. Theoretical pulse shapes. *a*—for trapping unimportant, *b*—uniform distribution of traps, and *c*—showing the time t_1.

time constant, a corresponding voltage pulse $v=q/C$ is produced at the amplifier input. Assuming direct recombination unimportant,[15] the fate of the secondary electrons is that they either become trapped within the volume of the crystal or that they are collected at the top electrode. Holes, on the other hand, remain localized near the region of ionization, or if they have mobility, drift a short distance to the lower electrode. Reversal of the battery polarity making the top electrode negative allows one to look for pulses of opposite sign due to the drift of holes.

The time dependence of the charge pulse, q, is derived making the following simplifying assumptions: (1) that n_0 electron-hole pairs are all released instantaneously at time $t=0$ at the bottom electrode, $x=0$; (2) that a uniform electric field, $E=V_c/d$ exists in the insulating crystal where d is the sample thickness; and (3) that there is a uniform density of electron traps in the volume of the crystal giving rise to an average time, T, before trapping. The average distance the conduction electrons drift before trapping is called the *Schubweg* (displacement distance), w, and is given by $w=\mu ET$ where μ is the electron mobility (assumed a constant). Processes such as thermal release from traps are not considered at this time. The number of electrons still free to move a time t after release of the ionization is

$$n=n_0e^{-t/T}. \tag{1}$$

The amount of charge induced on the electrodes by the drift of n electrons a small distance dx is

$$dq=ne\,dx/d. \tag{2}$$

With the aid of the substitution $dx=\mu E dt$, the charge $q(t)$ is found by integrating dq from 0 to time t.

$$q(t)=\frac{n_0e\mu E}{d}\int_0^t e^{-t/T}dt. \tag{3}$$

[15] The AgCl counter has been shown to have pulse-height response proportional to energy loss for particles ionizing near minimum, up to slow heavily-ionizing particles such as mesons stopping in the crystal. See H. G. Voorhies and J. C. Street, Phys. Rev. **76**, 1100 (1949). This proportionality fails for 5-Mev alpha particles, but this is believed due to the small penetration of the particles and a surface defect.

Expressing this in terms of w, we have

$$q(t)=(n_0ew/d)(1-e^{-\mu Et/w}), \tag{4}$$

which is applicable for times $t\leq d/\mu E$. For times greater than this the total charge induced is:

$$Q=(n_0ew/d)(1-e^{-d/w}). \tag{5}$$

A formula of this form was first derived by Hecht,[16] who also showed that it could be verified experimentally for electrons released by light. The effect of finite penetration of the beta rays can be taken into account by subtracting the average penetration from d in the exponent of the exponential. For the case that trapping is less important and w is very large so that $\mu Et\ll w$, Eq. (4) becomes

$$q(t)\simeq(n_0e\mu E/d)t. \tag{6}$$

Here the charge pulse has an approximately linear rise until a time $t_d=d/\mu E$ corresponding to collection of the n_0 electrons. In theory, a measurement of the time, t_d, allows one to compute the drift mobility from the relation

$$\mu=d/Et_d. \tag{7}$$

We shall show that this is also feasible in practice. It should be noted that even in the case that electron trapping is important, a break or discontinuity occurs at time t_d provided only that a fraction of the electrons reach the top electrode. A sketch of theoretical pulse shape is shown in Fig. 2. If one has an amplifier rise-time sufficiently short to make the detail of the pulse observable, all that is required to measure μ is a w long enough to approach saturation (collection of the majority of the secondary electrons). Mobility experimentally determined with the aid of Eq. (7) should be constant and independent of electric field strength; this condition, together with the observed pulse shape, serves as a check on the theory and its assumptions.

In the case that electrons are permanently trapped, a study of the shape of the pulses from an unsaturated crystal allows one to estimate the Schubweg, w. One convenient measure of the pulse shape is in terms of the quantity t_1 which is found from the intersection of the initial slope and the extrapolated final pulse height [see Fig. 2 (c)]. Defining t_1 in this way, $Q=q'(0)t_1$, where $q'(0)$ is the slope of the pulse at time $t=0$. The approximation used in Eq. (6) holds for the very beginning of the rise, so $q'(0)=n_0e\mu E/d$. Multiplying this by t_1 and using Eq. (5) we arrive at an expression independent of n_0:

$$w(1-e^{-d/w})=\mu Et_1. \tag{8}$$

Values of w obtained from an approximate solution of this equation can be compared with those obtained by the usual technique of fitting pulse height *versus* electric field data to the theoretical curve given by Hecht.[16]

[16] K. Hecht, Z. Physik **77**, 235 (1932).

The above analysis could not be applied if temporary trapping of secondary electrons were important. This would be indicated by unusually long pulse rises and low values of mobility as computed from Eq. (7). In this case, the observed collection time, t_0, would be more difficult to measure but it could be taken as the time for which the pulse approaches within a small fraction of its maximum value. It is given approximately by

$$t_0=(d/\mu E)+t_t, \tag{9}$$

where t_t is the average time electrons spend in shallow traps and μ is the mobility of an electron in the conduction band. We shall return to this point in Sec. IV B.

III. EXPERIMENTAL APPARATUS

A. Crystal Growth and Sample Preparation

As is evident from the preceding discussion, it was desired to work with crystals in which the electron range, w, at low temperature is at least as great as the crystal thickness, about 0.5 cm. This implies well-annealed, relatively perfect crystals and high field strengths. It was discovered in earlier experience with silver chloride[17,18] that crystals meeting these qualifications were not commercially available. Consequently the samples for the present experiments were grown from the melt[19–21] by taking advantage of the purification which takes place during crystallization and allowing the history of each sample tested to be known. Raw material was cut from large pieces of unrolled crystal supplied by the Harshaw Chemical Company. This was melted in air at temperatures not exceeding 25° above the melting point in cylindrical, flat-bottom, platinum-lined crucibles[22] which measured 1.5 cm in diameter and 7.6 cm in length. A crystal was then formed by slowly lowering the crucible and its contents through a steep temperature gradient in a furnace carefully stabilized by means of proportional control regulators. The rate of lowering was 1 mm per hour, and following solidification, the furnace and its contents were slowly reduced to room temperature over a period of 24 hours. The crystals could then be melted out by inserting the inverted crucible into a furnace set a few degrees above the melting point.

Disks of the order of 0.5 cm thick were cut from the crystal boules with a circular milling saw. Their surfaces were ground smooth and etched in sodium thiosulphate to remove metal or abrasive particles. Before the counting tests the samples were heat-treated after the manner of Haynes.[23] For this purpose we have found it satisfactory to rest the crystals on special reagent-grade MgO within a steel-lined copper box in the heat-treating furnace. After annealing, the crystals were handled very carefully to avoid the introduction of strains due to thermal gradients or cold working, and all manipulation was carried out in the dark or under red safelight. Thin high-conductivity electrodes were applied to the parallel surfaces of the crystals using silver-conducting paint.[24]

B. Optical Properties of the Crystals

Inspection under red polarized light indicated that the crystals were single and unstrained. Absorption of light in the tail of the fundamental band was compared for several samples after they had been used in the mobility experiments. The curve drawn through the open circles of Fig. 3 is typical of the absorption data taken with a Beckman Model DU spectrophotometer. The contribution to optical density due to successive reflections at the surfaces, but neglecting absorption of the reflected light, is shown as a dashed line in this figure. This reflection loss was computed by using the indices of refraction for AgCl given by Tilton *et al.*[25] Little variation in optical absorption was found between the four different crystals tested, including one sample cut directly from a large Harshaw crystal in which the Schubweg for electrons was of the order of one-tenth that in the recrystallized samples. It was not surprising to find the electronic properties much more sensitive to differences in composition or history of the crystals.

It should be emphasized that the low values of optical density for the long wavelengths shown in Fig. 3 (curve with open circles) were obtained only if the surfaces were very well polished and the crystal previously unexposed to blue light. Previous exposure or

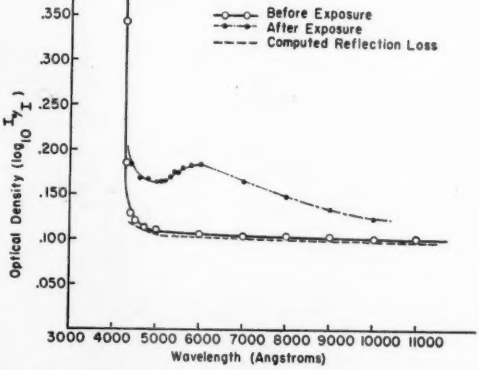


FIG. 3. Absorption of light in AgCl crystal II_2, thickness 0.65 cm, shown by the lower curve drawn through the open circles. The upper curve (solid dots) shows the darkening produced by long exposure to light of 4000 A.

[17] F. C. Brown and J. C. Street, Phys. Rev. **84**, 1183 (1951).
[18] W. L. Whittemore and J. C. Street, Phys. Rev. **76**, 1786 (1949).
[19] P. W. Bridgman, Proc. Am. Acad. Arts Sci. **60**, 305 (1925).
[20] D. C. Stockbarger, Rev. Sci. Instr. **7**, 133 (1936).
[21] R. H. McFee, J. Chem. Phys. **15**, 856 (1947).
[22] Obtained from Metals and Controls Corporation, General Plate Division, Attleboro, Massachusetts.
[23] J. R. Haynes, Rev. Sci. Instr. **19**, 51 (1948).
[24] Microcircuits Corporation, New Buffalo, New York.
[25] Tilton, Plyler, and Stephens, J. Opt. Soc. Am. **40**, 540 (1950).

surface properties may explain the slightly different results obtained by Gilleo[26] in the long-wavelength region. The effect of exposure of our crystals at room temperature to light of 4000 A from the spectrophotometer for 10 minutes is shown by the solid dots. The crystals were not observed to darken very rapidly and there was evidence for saturation of the darkening in the molecular range as observed by Löhle.[27]

C. Crystal Counter Apparatus

For the mobility measurements the crystals were mounted in the cryostat outlined in Fig. 4. The crystal rests on a brass shelf which is supported by a stainless steel tube forming the inner section of a Dewar. To slow down the rate of cooling, and to provide for intermediate temperature, the liquid nitrogen was sometimes poured into an inner cylindrical reservoir (not shown) which was supported and insulated by a piece of styrofoam. To aid in maintaining thermal equilibrium between the crystal and shelf, the closed space immediately surrounding the crystal was filled with dry nitrogen gas under pressure. A small-diameter thin-wall stainless steel tube leads from this space to the outside for evacuation and filling. Contact between crystal and shelf when in vacuum was not found adequate in the earlier runs, even though a drop of Silicone oil was placed at the interface. One of the difficulties here was that, because of the introduction of strains, the top electrode could bear only very lightly on the crystal. So that the crystal temperature could be related to the shelf temperature, a small copper-constantan thermocouple was imbedded in a hole in a test crystal, and measurements were taken. Vacuum pumps maintained a pressure of less than 10^{-4} mm of mercury in the space surrounding the shelf and liquid nitrogen reservoir.

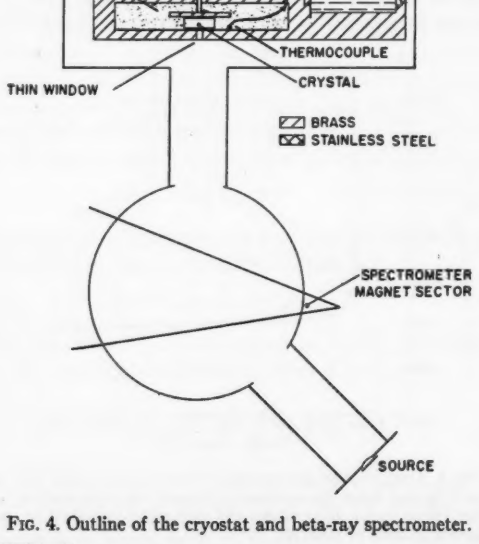


FIG. 4. Outline of the cryostat and beta-ray spectrometer.

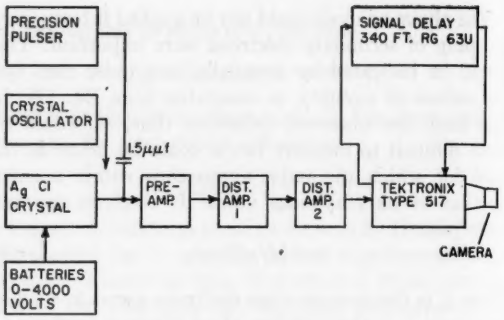


FIG. 5. Block diagram of electronic circuits.

A simple beta-ray spectrometer utilizing wedge-shaped magnet pole pieces and incorporating double focusing[28] was used as the source of ionizing particles. Such a design had the advantage of simplicity, ease of construction, good shielding against gamma radiation, and the location of the source and detector (the crystal) outside of the magnetic field. It had the disadvantages of relatively small solid angle and limited resolution. A Ru^{106}-Rh^{106} source allowed the selection of roughly monoenergetic electrons from a few tenths of a Mev to over 3.0 Mev. Calibration at one energy was achieved by means of the internal conversion line of Cs^{137}.[29] For this purpose, and also to observe the counting efficiency of the crystals, a thin-window Geiger counter could be installed in place of the crystal above the hole in the brass shelf.

A block diagram of the crystal-counter circuits is shown in Fig. 5. A single beta particle releases only about 10^{-14} coulombs of charge in the crystal, and it is required that the amplifier faithfully pass the detail of a pulse which has a rise-time of as little as 0.10 μsec. With the exception of the preamplifier, distributed amplification was employed throughout. A low-noise preamplifier circuit was developed which matched the high impedance of the crystal to the 200-ohm impedance of the distributed amplifiers. Its first stage employed a cascode-connected 6BQ7, which proved to be more stable, and to have lower input capacity than the more conventional arrangement. An over-all rise-time was attained of 0.018 μsec. The clipping time of the system was greater than 10 microseconds so that slower rising pulses could be studied, although usually in this case

[26] M. A. Gilleo, Phys. Rev. **91**, 534 (1953).
[27] F. Löhle, Göttingen Nachr. **2**, 271 (1933).
[28] W. G. Cross, Rev. Sci. Instr. **22**, 717 (1951).
[29] L. M. Langer and R. D. Moffat, Phys. Rev. **78**, 74 (1950).

the bandwidth was reduced with an improvement in signal to noise ratio. Amplitude calibration was achieved by applying a known amount of charge at the preamplifier input through a 1.5-$\mu\mu$f series capacitor. The timing calibration of the Tektronix-type 517 oscilloscope was checked with a crystal oscillator. Oscilloscope traces were photographed on 35 mm film and the pulses analyzed after projection.

IV. RESULTS AND DISCUSSION

A. Electron Drift Mobility

Figure 6 shows a group of pulses due to 2.0-Mev beta particles incident upon crystal number V_1 when the crystal temperature was 105°K. The pulses illustrate the general conditions under which the mobility data were taken for all the samples reported in Table I. Amplifier rise-time was of the order of 0.018 μsec and the voltage gain was adjusted so that 1 cm of pulsed height corresponded to 2×10^{-14} coulombs of charge at the crystal. For the first group of pulses on the left, the field was 9755 volts/cm and the sweep-time 0.110 μsec per cm. The essential features of the simple theory outlined in Sec. II are verified. An approximately linear rise indicates that the crystal was operating under nearly saturation conditions, as was verified also by little change in pulse height at different field strengths. Pulses with a lower field of 2250 volts/cm are shown below the test pulse on the right. Amplifier gain was the same as before but sweep-time was changed to 0.525 μsec/cm.

From ionization energy-loss considerations a 2.0-Mev electron should penetrate AgCl a distance of about 2.0 mm; a 0.5-Mev electron should penetrate only about 0.5 mm. However, the average penetration is less than this due to scattering of the beta particle, and the important quantity for purposes of computing pulse height is the center of gravity of the ionization as measured from the bottom electrode. For 3.0-Mev electrons the pulses became slightly rounded near the top of the rise, but less effect of this sort was noticed at lower energies where the rise-times were measured.

The results of a study of a large number of pulses for various energies selected by the spectrometer are in agreement with those of Van Heerden,[8] who obtained approximate linearity of pulse height with energy for beta rays. For less penetrating radiation such as alpha particles, there seemed to exist a defective surface layer which sometimes could be observed for low-energy beta particles. When this condition existed, strains were visible under polarized light near the lower electrode. The condition could be improved by reannealing and then slow cooling to low temperature with less pressure of the top electrode on the crystal. In a properly cooled crystal (taking four hours or longer to reach −75°C), it was possible to demonstrate that these defective layers were at least less than 0.2 mm.

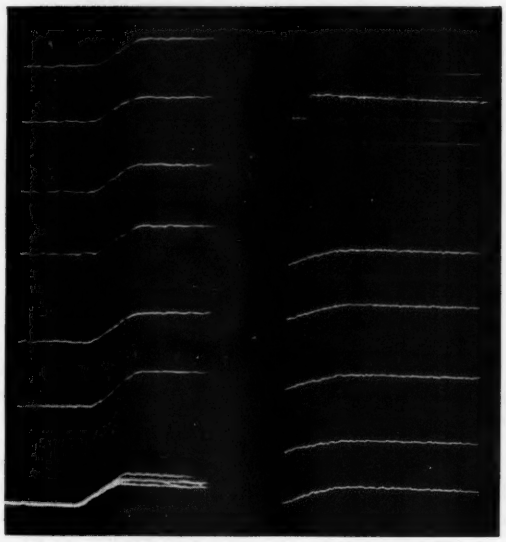


FIG. 6. Examples of crystal counter pulses produced by 2.0-Mev beta particles. At the left the field $E=9755$ volts/cm, the sweep time, 0.110 μsec/cm. A graticule ruled with 1-cm squares is shown on the right and superimposed is a test pulse of 2×10^{-14} coulomb. Below on the right $E=2250$ volts/cm and sweep-time, 0.525 μsec/cm.

TABLE I. Electron mobility and range in annealed AgCl at low temperature. Sample H was cut directly from stock material without recrystallization.

Sample	Thickness d (cm)	w/E (cm²/volt)	Temperature (°K)	Mobility, $\bar{\mu}$ (cm²/volt sec)
V_1	0.408	8.2×10^{-5}	86±3	274±13
IV_1	0.450	1.1×10^{-4}	86	267
IV_2	0.462	1.4×10^{-4}	86	300
II_1	0.650	2×10^{-4}	86	268
H	0.570	1×10^{-5}	130	

Electron mobility computed with the aid of Eq. (8) is tabulated for several crystals in Table I. The following procedure was used in arriving at the values of $\bar{\mu}$ given in the last column. Measurements were made of the time, t_d, for 50 to 100 pulses recorded at each condition of electric field and temperature. At high fields it was found that there was little scattering of the values of t_d from pulse to pulse. For example, in the most studied crystal, number V_1, the value of the average of t_d for a group of data was 0.149 μsec at 86°K and 9755 volts/cm, whereas the root-mean-square deviation in t_d was 0.009 μsec or about 6 percent. At high fields the deviation was about what would be expected due to random noise on the pulse; at lower voltages, more variation occurred. Using the average of t_d for a group of pulses, mobility was computed from Eq. (8) and the results plotted as a function of electric field. Examples of the dependence of μ on E for crystal V_1 are shown in Fig. 7. It is seen that the values of μ become constant at high fields. Finally, the mobility $\bar{\mu}$ listed in Table I are the means of these nearly constant high-field values.

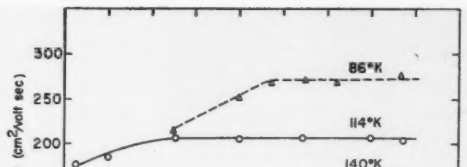


FIG. 7. Dependence of μ on E for crystal V_1. The data for 86°K shows the effect of strains introduced by too rapid cooling from 114°K.

Crystal number V_1 was the most carefully studied. The error in $\bar{\mu}$ was estimated from the deviations in t_d and the other quantities to be less than ± 5 percent. The temperature in this case was known to within ± 3°K. Actually three separate runs were made on this crystal (each after annealing) with good reproducibility. Mobilities for the other crystals were more approximate and the temperatures less accurately determined. However, again some of the crystals were tested several times with consistent results.

The quantity w/E listed in Table I was estimated by fitting saturation curves of the type given by Eq. (6) to pulse height *versus* field data. In some cases where deep traps were predominant, a crude check could be had from the pulse shapes using Eq. (9). For the good crystals, w at a given field was not a constant for different runs, but depended upon the conditions of strain of the crystal.

Although the pulses at high field have a linear rise in agreement with theory, two objections may be brought up concerning the use of Eq. (5) and the simple theory from which it was derived in computing mobility: (1) It might be stated that the existence of a volume distribution of charge such as produces polarization of a crystal counter may cause the internal field to be different than given by $E=V_e/d$; and (2) the distance the secondary electrons move may not be d, the thickness of the crystal, due to boundaries or trapping layers in the volume of the crystal. That polarization effects were not important in our case can be deduced from the following arguments. Silver chloride at room temperature is an ionic conductor which, in the absence of an externally applied electric field, should be charge neutral throughout its volume. The crystals were cooled with shorted electrodes to low temperature, where the ionic conductivity is negligible. The collecting voltage was then applied and *single particles* detected at low counting rates, usually of the order of 100 counts/minute. Any variation of the pulse height could be observed as the experiment progressed. Moreover, using a gamma ray source, a test could be made for the existence of internal fields due to bound charge by removing V_e and looking for pulses of opposite polarity. Polarization effects did not begin to become important until about 10^4 to 10^5 counts, which was more than registered during the runs. This is in agreement with the results reported by others for AgCl.[30] We have here one of the chief advantages of the method as compared with bombardment-induced conductivity techniques in which polarization effects are certainly more important.[31]

That trapping layers were not important is shown by the consistent values of mobility obtained from crystal to crystal, and for a crystal operated first in one position and then inverted. In addition, early tests were made in which the crystal could be moved around on the brass shelf without finding very marked variations over the crystal area. An important consideration bearing on these matters is the average energy for production of one electron-hole pair, a quantity which we will call ϵ. For the mobility runs, ϵ was of the order or less than 10 electron volts per ion pair. It was carefully measured for crystal V_1 by setting the spectrometer to the internal conversion line of Cs^{137}. A pulse-height distribution for a total of 550 pulses is shown in Fig. 8. This is in agreement with previous results[8] where it was argued that the distribution of smaller pulses was primarily due to the scattering of the beta particles. The peak of the distribution corresponds to 10^{-14} coulombs of induced charge. After making small corrections for window thickness, lack of saturation of the crystal and particle penetration one gets a value of $\epsilon=7.5\pm0.5$ electron volts/ion pair.

The dependence of drift mobility on temperature is shown in Fig. 9, which summarizes data for three crystals and shows also the higher-temperature points given by Haynes and Shockley.[7] The mobility measurements of Fig. 9 are accurate to within about 10 percent for crystals II, and IV, and within 5 percent for

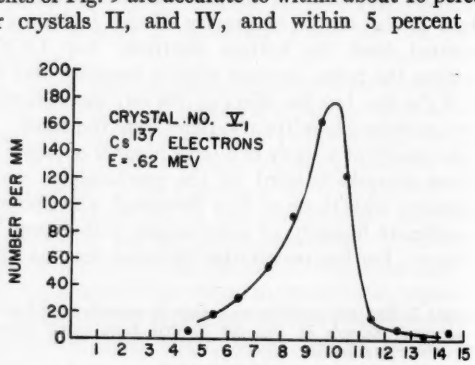


FIG. 8. Pulse height distribution for monoenergetic, 0.62-Mev beta rays. A maximum near 10 mm corresponds to 10^{-14} coulombs and yields a value of $\epsilon=7.5\pm0.5$ ev/ion pair.

[30] L. F. Wouters and R. S. Christian, U. S. Atomic Energy Commission document MDDC-1324, 1947 (unpublished).

[31] K. G. McKay, Phys. Rev. **74**, 1606 (1948); and **77**, 816 (1950).

crystal V, with an uncertainty in temperature of, at most, ±5°K. Data for only three crystals are given, but it is believed that the results are representative of μ *versus* T for AgCl. The values at low temperature are in agreement with the early crystal counter work on AgCl; for instance, Hofstadter cites values from 114 to 400 cm²/volt sec at 77°K. Moreover, in view of the constancy of μ *vs* E at high fields, it is reasonable to assume that the drift mobility we have measured is close to the so-called microscopic mobility in AgCl.[7] In this connection note that Hall mobility[13] measured from the Hall angle (not corrected by $8/3\pi$) in one of our samples at 86°K lies just above the drift mobility.

Let us first compare the data with the theory of Low and Pines,[3] who assume that electrons are scattered primarily by optical modes of vibration, as was also assumed in the earlier work of Fröhlich and Mott.[4] They give an expression which is applicable only for $T \ll \Theta$, as follows:

$$\mu = \frac{1}{2\alpha\omega}\frac{e}{m}\left(\frac{m}{m^*}\right)^2 f(\alpha)(e^{\Theta/T}-1), \qquad (10)$$

where we have subtracted 1 from the exponential to give the proper convergence at higher temperatures. The constant α measures the strength of the electron-lattice interaction and is proportional to $m^{\frac{1}{2}}$; $f(\alpha)$ is a slowly varying function about 1.14 in our case; m is the effective mass of the electron in the conduction band and, since this is unknown, plays the role of an adjustable parameter in the theory; and finally, the effective mass of electron plus associated lattice polarization is given by $m^*=m(1+\alpha/6)$. The Debye temperature $\Theta=\hbar\omega/k$ corresponds to the frequency, ω, of the longitudinal optical modes of vibration of the lattice. Using $\omega=\omega_t(\epsilon/n^2)^{\frac{1}{2}}$, we can deduce Θ from the low- and high-frequency dielectric constants, $\epsilon=12.3$, $n^2=4.04$,[32] and the reststrahl wavelength $\lambda_t=90$ microns.[33] We arrive at $\Theta=280$°K for AgCl. Now if the experimental points for μ below about 140°K are plotted on a semilog basis against $1/T$, it is quite clear that they do *not* lie along an $e^{280/T}-1$ curve (rough agreement can be had by assuming an unreasonably low Debye temperature $\Theta \simeq 100$°K).

On the other hand, the data are seen to agree quite well over a wide temperature range with a $T^{-\frac{3}{2}}$ law as would be predicted by scattering with the acoustical modes of vibration.[34] The solid line of Fig. 9 is given by

$$\mu = 2.54\times10^5 T^{-\frac{3}{2}}, \qquad (11)$$

which was shown by Haynes and Shockley to fit their results. If, in order to match the $T^{-\frac{3}{2}}$ data, we assume that the interaction over the entire temperature range

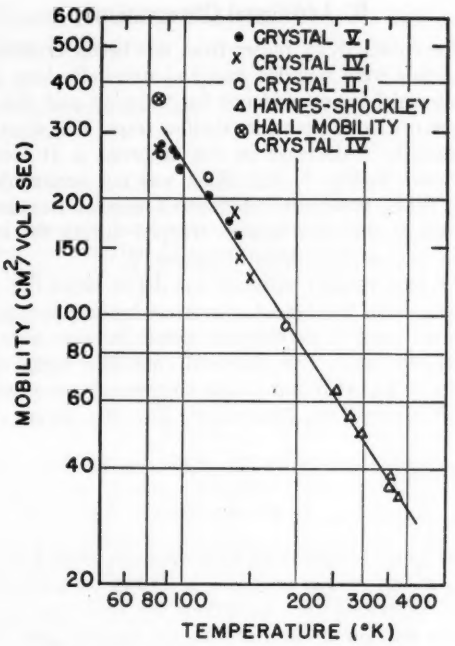


FIG. 9. Mobility as a function of temperature for electrons in AgCl.

is primarily acoustic, and, at the same time, that the theory based on optical modes is applicable below say 140°K, it is necessary to pick the coefficient of $e^{\Theta/T}-1$ in Eq. 10 to be about 50 or larger. This corresponds to an effective mass $m \simeq 0.2m_e$, where m_e is the free electron mass. Now, although the applicability of an acoustic equation, such as Wilson[34] derives, to an ionic crystal is not clear, we can roughly estimate mobility for $m \simeq 0.2m_e$ and arrive at values which are quite a bit higher than the experimental points. It seems necessary to pick a larger effective mass for agreement with the acoustic theory.

There is an unresolved difficulty here. The theoretical situation, which is exceedingly complex, merits further examination. On the experimental side, it would be interesting to have more data at lower temperatures. It is conceivable that scattering by dislocations or other imperfections is important at 86°K, but in this case one would expect greater variation in mobility from sample to sample. A theory based on ionized impurity scattering along with optical scattering might explain the observed results, but a fairly high concentration of impurity atoms is required.[35a]

[32] K. Hojendahl, Kgl. Danske Videnskab. Selskab, Mat.-fys. Medd. **16**, No. 2 (1938).

[33] R. B. Barnes, Z. Physik **75**, 723 (1932), extrapolating for AgCl between the points given.

[34] A. H. Wilson, *The Theory of Metals* (Cambridge University Press, London, 1953), p. 265.

[35] The work of C. Allemand and J. Rossel, Helv. Phys. Acta. **27**, 212 (1954), has recently come to the attention of the author. It would appear that the composition of their crystals is quite different from ours.

[35a] *Note added in proof.*—In a private communication D. Pines has pointed out an error in the expression for mobility given by reference 3. Equation 10 should contain the factor m/m^* raised to the third power rather than the second, that is, μ should be reduced by a factor m/m^*. This makes little change, however, in the above arguments.

B. Additional Observations

The conductivity pulses from our better crystals at low fields were rounded and had unusually long rise-times which gives evidence for trapping and delayed release of charge from very shallow traps. This situation resulted in a decrease in the observed μ at low E, as shown in Fig. 7. The effect was not noticeable at high fields, presumably because a smaller fraction of secondary electrons became trapped during the short times to cross the crystal thickness, d.

We may roughly estimate the depth below the conduction band for these shallow traps by first finding the average time, t_t, an electron spends in traps near the lower electrode. The observed collection time, t_0, is given by Eq. (10) and to this corresponds an observed mobility $\mu_0 = d/Et_0$. Rearranging Eq. (10), we get

$$\frac{1}{\mu_0} = \frac{1}{\mu} + \frac{t_t E}{d}. \tag{12}$$

Thus t_t may be estimated from the slope of the $1/\mu_0$ *vs* E data at low E. For crystal V at 140°K we find $t_t \simeq 0.5$ μsec. Supposing that conditions are such that a negligible fraction of the electrons are trapped more than once, it is reasonable to associate this t_t with the average time an electron spends in a trap. Saying that the probability per unit time for escape from a trap $B = B_0 e^{-E/kT} = 2\times 10^6$ sec^{-1}, the thermal trap depth, E, is about 0.1 electron volts, depending slightly upon what value of B_0 is chosen.[36] Traps of about this depth have previously been found in AgCl from the results of "glow curve" experiments.[37]

It is quite clear that strains play an important role in affecting the range of electrons in AgCl. The concentration of both the shallow, 0.1 ev, traps and of deeper traps could be increased by too rapid cooling or by mild cold-working (such as produced by moving the crystal as it rested under the top electrode on the brass shelf). There was evidence that these effects depended somewhat upon temperature; for example, a crystal was more resistant to the effects of strains while at the lowest temperature 86°K. Figure 7 shows how shallow traps became more important (μ dropping off at a higher E) in crystal V after rapidly cooling from 114°K to 86°K. Subsequent examination under red polarized light showed the presence of strains near the bottom electrode which could be removed as usual by heat treatment.

Simpson[38] has calculated the thermal trap depth of an interstitial silver ion in AgCl and arrived at about 0.1 electron volts. It is possible that cold-working increases the concentration of interstitial silver in our crystals. Such would be the case if Frenkel defects were generated in sufficient numbers by dislocations during cold work.[39]

A search was made in all of the crystals for evidence that positive holes drift with ranges comparable to that of electrons. Hole pulses were not observed at temperatures from 86°K to 150°K. Small pulses of opposite polarity were observed under reversed field (top electrode negative), but only for the most penetrating beta particles. They could be explained in terms of sweeping of the secondary electrons toward the lower electrode. We are led to conclude that if holes have a mobility comparable with electrons, their range is less than $\frac{1}{10}$ that of electrons. We would not have been able to detect the migration of holes with very low drift velocity.

The author would like to express his appreciation for the interest and assistance of Professors J. C. Street, R. J. Maurer, W. L. Parker, and R. T. Ellickson. He would also like to thank Mr. H. Vollum and Mr. C. Moulton of Tektronix, Inc., for their help. It is a pleasure to acknowledge discussions with Professor J. Bardeen, D. Pines, F. Seitz, and Dr. R. T. Thomson. The beta-ray spectrometer was the work of Mr. M. Crooks, and Mr. W. Sandman prepared one of the best crystals.

[36] N. F. Mott and R. W. Gurney, *Electronic Processes in Ionic Crystals* (Oxford University Press, London, 1940), p. 108.
[37] A. Moore, thesis, Cornell University, 1951 (unpublished).
[38] J. H. Simpson, Proc. Roy. Soc. (London) **A197**, 269 (1949).
[39] F. Seitz, Advances in Physics **1**, 43 (1952).

Vibration Spectra and Specific Heats of Cubic Metals. I. Theory and Application to Sodium

A. B. BHATIA*
Division of Physics, National Research Council, Ottawa, Canada
(Received October 14, 1954)

A model for monovalent metals is proposed, on the basis of which the secular equations determining the frequencies of the normal modes of vibration are derived for body-centered and face-centered cubic metals. In contrast to the two force constants occurring in the usual treatments based on the Born-von Kármán model, each of these equations—one for each type of lattice—contains three force constants which may be identified with the three independent elastic constants of a cubic metal. The limitations of the model are discussed.

The frequency spectrum of sodium, and hence its specific heat C_v at constant volume, are calculated. It is found that the agreement between the observed and calculated values of C_v is more satisfactory in the entire relevant temperature range, and especially at low temperatures, than that obtained by Bauer using the customary two force-constant model.

I. INTRODUCTION

ALTHOUGH a proper theory for calculating the frequencies ω of the normal modes of vibration of an ionic or homovalent crystal exists and has been applied by several authors,[1] the corresponding situation with regard to metals is not so satisfactory. This is because, whereas in the former types of crystals the forces on the ions (atoms) can be described in terms of an interaction potential function between the ions (atoms), in the latter the interaction between the ions and the valence electrons is of decisive importance. Hence, in general, the forces on ions in a metal are not describable in terms of an interaction potential function between the ions alone, and this makes any rigorous treatment of the problem difficult. It is customary, therefore, to calculate the frequencies of normal modes of vibration for metals on the basis of the Born-von Kármán[2] model. One considers the interactions to be significant only between the nearest and the next nearest neighbors and assumes Hooke's law of forces. The classical equations of motion for the ions are then set up and solved with the aid of Born's cyclic boundary conditions. These equations of motion naturally contain two Hooke's constants, usually denoted in literature by α and γ, which respectively correspond to nearest neighbor and next nearest neighbor interactions. The force constants α and γ are determined from the elastic constants by comparing expressions (in terms of α and γ) for the velocities of long elastic waves with the corresponding expressions of the elasticity theory. Now for cubic crystals there are three elastic constants c_{11}, c_{12}, and c_{44} from which α and γ have to be determined. This is possible only if there is a relation between the three elastic constants. One relation between them exists either if the crystal is elastically isotropic $\epsilon=c_{11}-c_{12}-2c_{44}=0$, or if the *entire* interactions are central when $c_{12}=c_{44}$ (Cauchy's relation). For the alkali and nobel metals with which we shall be primarily concerned in the present work, the isotropy condition is far from satisfied; the quantity $2c_{44}(c_{11}-c_{12})^{-1}$, which is unity for an elastically isotropic crystal, is of the order 8 and 3 for the alkali and noble metals respectively.

The Cauchy relation ($c_{12}=c_{44}$) also does not hold for metals. This is because certain energies—like the energy E_0 of the ground state of the valence electron and the Fermi, exchange, and correlation energies E_1 of the conduction electrons—depend on the atomic volume alone,[3] so that the forces on the displaced ions arising from these energies cannot be regarded as due to a central interaction potential between the ions. In addition to the forces arising from E_0 and E_1, there are contributions also from the exchange repulsion, van der Waals forces, ion-ion (screened) Coulomb interaction, etc.[4]; to a first approximation, all these latter may be described by a central interaction between the ions. We denote the corresponding potential function by $W(r)$, where r is the distance between the ions. If we distinguish the contributions to elastic constants from $W(r)$ and Es' by superscripts W and E respectively, we may show, following Fuchs,[3] that the contributions to combinations of elastic constants corresponding to shearing strains (no change in volume) come only from W (and not from E_0 and E_1). Thus

$$c_{11}{}^W-c_{12}{}^W=c_{11}-c_{12}, \quad c_{44}{}^W=c_{44},$$

and

$$c_{44}{}^E=c_{11}{}^E-c_{12}{}^E=0. \tag{1}$$

The Born-von Kármán model has been applied to metals by Fine,[5] Leighton,[6] and Bauer.[7] Fine calculated

* National Research Laboratories Postdoctorate Fellow.

[1] For an excellent account see M. Born and K. Huang, *Dynamical Theory of Crystal Lattices* (Oxford University Press, England, 1954).

[2] M. Born and T. von Kármán, Physik. Z. **13**, 297 (1912); **14**, 15 (1913).

[3] K. Fuchs, Proc. Roy. Soc. (London) **A153**, 622 (1935); **A157**, 444 (1936).

[4] See, for example, N. F. Mott and H. Jones, *The Theory of the Properties of Metals and Alloys* (Oxford University Press, England, 1936), p. 148.

[5] P. C. Fine, Phys. Rev. **56**, 355 (1939).

[6] R. B. Leighton, Revs. Modern Phys. **20**, 165 (1948).

[7] E. Bauer, Phys. Rev. **92**, 58 (1953).

the frequency spectrum for tungsten which is elastically isotropic so that, in this case, there are only two independent elastic constants. On the other hand, Leighton and Bauer, whose calculations refer to silver and sodium respectively, essentially assume that c_{11}^E and c_{12}^E do not contribute to the motion of the ions and that $c_{44}^W = c_{12}^W$, thus having only two independent elastic constants c_{11}^W and $c_{12}^W = c_{44}^W$ in their models. It should be noted that even though c_{12}^W and c_{44}^W arise wholly from a central interaction, these two constants need not necessarily be equal since the Cauchy relation $c_{12} = c_{44}$ holds only if the entire interactions under which the cubic crystal is in equilibrium are central. We should mention that both these authors were aware of the limitations of their models.

From the foregoing it is clear that for cubic metals, which are neither isotropic nor satisfy the Cauchy relation, one should set up equations of motion for the ions which contain three force constants and which, for the limiting case of infinitely long waves, reduce to the corresponding equations of the elasticity theory. In this paper we consider a *model* for a metal which enables one to set up such equations.[8] It is assumed[8a] that the forces on an ion may be regarded as arising from (a) a central interaction $W(r)$ between the ions which is significant between the nearest neighbors only and (b) from certain energies E which depend on the *atomic volume* only and which are due to the presence of the free electrons and their interactions with the ions. In order to calculate the effect of (b) on ionic vibrations it is assumed, following the Sommerfeld model, that in the absence of thermal motion the ionic charge it uniformly smeared out over the entire metal. The electrostatic potential φ of the system is then a constant ($=\varphi_0$, say). The change in φ due to a given displacement of the ions is calculated by the Thomas-Fermi-Dirac method. The force on an ion is then just $-e\,\mathrm{grad}\varphi$. Next the classical equations of motion are set up and solved for the body-centered and face-centered cubic lattices. Each of the resulting secular equations in ω^2—one for each type of lattice—contains three force constants which are identified with the three elastic constants of a cubic crystal. In each case, the secular equation for the limiting case of infinitely long waves then reduces to the well-known Christoffel equations of the elasticity theory.

Clearly, the assumptions and the consequent limitations underlying the procedure adopted here need some discussion. For convenience, this is deferred until the secular equations have been derived.

In Sec. III, the frequency spectrum of sodium, and hence its specific heat C_v at constant volume as a function of temperature, are calculated. It is found that the agreement between the calculated and observed values of C_v is more satisfactory than that obtained by Bauer[7] using the customary two force-constant model.

Similar calculations for silver which has face-centered cubic structure will be given in Part II of this work.

II. THE SECULAR EQUATION CALCULATION OF φ FOR A GIVEN DISPLACEMENT OF THE IONS

As mentioned in the introduction, it will be assumed for the purposes of this calculation that the ionic charge is smeared out over the metal. Then in the absence of thermal motion, the ionic and the electronic charge densities are equal and constant at every point of the metal and the electrostatic potential φ of the system is also a constant, equal to φ_0, say. When an elastic wave passes through a metal, the ionic charge density and hence the electronic charge density will vary from point to point. This will result in a change in φ which will depend now on space and time coordinates. Let $\varphi = \varphi_0 + \varphi_1(\mathbf{r})$. For small amplitudes of the elastic wave, we may take $\varphi_1(\mathbf{r}) \ll \varphi_0$. One then also has the result that the changes in φ are small over regions of linear dimensions of the order of the de Broglie wavelength of the Fermi electrons. This condition is necessary for the validity of the Thomas-Fermi method. Hence we may calculate φ_1 for a given ionic displacement either by the Thomas-Fermi method (without exchange) or by the Thomas-Fermi-Dirac method (with exchange). For convenience, the calculations will be first made by the former method.

Let E_m denote the energy of the highest level occupied by a conduction electron. We may take E_m to be unchanged when the metal is disturbed by the passage of an elastic wave since it will not change the *over-all* volume of the metal. Then in the Thomas-Fermi method the electronic number density $n(\mathbf{r})$ and the potential $\varphi(\mathbf{r})$ are related by the equation[9] (choosing the zero for $\varphi(\mathbf{r})$ such that $E_m = 0$):

$$n(\mathbf{r}) = (8\pi/3h^3)(2me)^{\frac{3}{2}}[\varphi(\mathbf{r})]^{\frac{3}{2}}, \tag{2}$$

where $-e$ and m are the electronic charge and mass respectively, and other symbols have their usual meaning. The average density n_0 of the electrons is obtained from (2) by putting $\varphi = \varphi_0$ in it. Then remembering that $\varphi_1 \ll \varphi_0$, we have

$$\Delta n(\mathbf{r}) = n(\mathbf{r}) - n_0 = \eta' \varphi_1(\mathbf{r}), \tag{3}$$

where

$$\eta' = (4\pi me/h^2)(3/\pi)^{\frac{1}{3}} \times n_0^{\frac{1}{3}}. \tag{4}$$

Now let $\mathbf{u}(\mathbf{r})$ denote the magnitude and direction of the displacement of the ion at $\mathbf{r}$ from its normal position;

[8] In the present work the discussion will be confined to the alkali and the noble metals which have body- and face-centered cubic structures respectively. To a good approximation, one may regard all the valence electrons to be free in both these groups of metals.

[8a] If one assumes a noncentral interaction and considers nearest and next nearest neighbour interactions only, one has five force constants in the scheme of Born and von Kármán. These can be determined in terms of the elastic constants only if there are two relations between the five force constants; thus it seems necessary to introduce explicit assumptions about the nature of interactions in a metal (see reference 5).

[9] See, for example, reference 4, pp. 48–49.

the change $\Delta N(\mathbf{r})$ in the ionic number density (average number $=N_0\equiv n_0$) is given by

$$\Delta N=N-N_0=-N_0 \operatorname{div} \mathbf{u}(\mathbf{r}). \tag{5}$$

Using (3) and (5) and applying Poisson's equation, we have

$$\nabla^2(\varphi_0+\varphi_1)=\nabla^2\varphi_1=-4\pi\rho=4\pi e[\eta'\varphi_1-\Delta N(\mathbf{r})]. \tag{6}$$

Since $\varphi_1(\mathbf{r})\equiv 0$ when $\Delta N(\mathbf{r})=0$, only the particular integral of (6) is of interest. Hence, formally,

$$\varphi_1(\mathbf{r})=-4\pi e\Delta N(\mathbf{r})/(\nabla^2-4\pi e\eta'). \tag{7}$$

The force on an ion at $\mathbf{r}$ (charge $+e$) due to $\varphi(=\varphi_0+\varphi_1)$ is given by

$$\mathbf{F}(\mathbf{r})=-e\operatorname{grad}\varphi=-e\operatorname{grad}\varphi_1(\mathbf{r}). \tag{8}$$

Let $\mathbf{u}(\mathbf{r})$ be of the form

$$\mathbf{u}(\mathbf{r})=\mathbf{e}Ae^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}, \tag{9}$$

where A is the amplitude, $\mathbf{e}$ the unit vector in the direction of the displacements, $\mathbf{k}$ the wave-vector ($|k|=2\pi/\lambda$), and ω is the angular frequency. Then $\mathbf{F}(\mathbf{r})$ is given by

$$\mathbf{F}(\mathbf{r})=-\mathbf{k}\frac{4\pi e^2N_0(\mathbf{e}\cdot\mathbf{k})}{4\pi e\eta'+k^2}A\exp[i(\mathbf{k}\cdot\mathbf{r}-\omega t)]. \tag{10}$$

The effect of exchange interaction between the electrons is easily included in the above calculation by Dirac's[10] modification of the Thomas-Fermi method, since one has only to replace (2) by the equation[11]

$$n(\mathbf{r})=(8\pi/3h^3)(2me)^{\frac{3}{2}}[(\varphi+\tau_0^2)^{\frac{1}{2}}+\tau_0]^3 \tag{2a}$$

with $\tau_0=(2me^3/h^2)^{\frac{1}{2}}$. The average electronic number density n_0 is now obtained by putting $\varphi=\varphi_0$ in (2a). Writing again $\varphi=\varphi_0+\varphi_1$ and remembering that $\varphi_1\ll\varphi_0$, it may be seen that, to include the exchange effects, one has to replace in (3) and all subsequent equations the constant η' by η, where

$$\begin{aligned}(\eta)^{-1}&=(\eta')^{-1}-\tfrac{1}{3}(3/\pi)^{\frac{1}{3}}en_0^{-\frac{1}{3}}\\&=(\eta')^{-1}-0.33en_0^{-\frac{1}{3}}.\end{aligned} \tag{11}$$

Derivation of the Secular Equation

Next we have to calculate the force on an ion due to the interaction potential $W(r)$ assumed to be effective only between the nearest neighbor ions in a lattice. Let the ion under consideration be situated on the lattice point at the origin of a Cartesian coordinate system and let the coordinates of the neighboring lattice points be $\mathbf{r}_{lmn}{}^{(0)}(la, ma, na)$. Here a is the distance between the nearest neighbors and (l, m, n) the direction cosines of the line joining the origin and a nearest neighbor. By making a suitable choice for the coordinate system, we may take (l, n, m) to be $|\sqrt{3}|^{-1}\times(\pm1, \pm1, \pm1)$ for body-centered cubic (b.c.c.) lattice which has eight nearest neighbors; similarly, for a face-centered cubic (f.c.c.) lattice which has twelve nearest neighbors, (l, m, n) may be taken as $|\sqrt{2}|^{-1}(\pm1, \pm1, 0)$, $|\sqrt{2}|^{-1}(\pm1, 0, \pm1)$ and $|\sqrt{2}|^{-1}(0, \pm1, \pm1)$. Denoting the displacements of the central and nearest neighbor ions from their normal positions by $\mathbf{u}_0$ and $\mathbf{u}_{lmn}$ respectively and denoting by V the contribution to the potential energy (due to W) of the crystal from the terms which involve the coordinates of the ion at the origin, we have

$$V=\sum_{l,m,n}W(|\mathbf{r}_{lmn}{}^{(0)}+\mathbf{u}_{lmn}-\mathbf{u}_0|).$$

Expanding the right-hand side by Taylor's theorem, we obtain

$$\begin{aligned}V=\sum_{l,m,n}\Bigg[&\left\{\frac{1}{r}\frac{dW}{dr}\right\}_{r=a}[\mathbf{r}_{lmn}{}^{(0)}\cdot(\mathbf{u}_{lmn}-\mathbf{u}_0)\\&+\tfrac{1}{2}|\mathbf{u}_{lmn}-\mathbf{u}_0|^2]+\frac{1}{2}\left\{\frac{1}{r}\frac{d}{dr}\left(\frac{1}{r}\frac{dW}{dr}\right)\right\}_{r=a}\\&\times[\mathbf{r}_{lmn}{}^{(0)}\cdot(\mathbf{u}_{lmn}-\mathbf{u}_0)]^2\Bigg]+O(u^3),\end{aligned} \tag{12}$$

where the summation is over all the nearest neighbors of the central ion. It should be noticed that $(dW/dr)_{r=a}$ is not zero in our case since $W(r)$ is not the entire potential which determines the equilibrium distance a. The force on a particle at the origin due to (12) is just $-\operatorname{grad}_{\mathbf{u}_0}V$.

Combining this last result with (10) we may write the equation of motion for the ion at the origin in the form

$$M(\partial^2/\partial t^2)\mathbf{u}_0=-\operatorname{grad}_{\mathbf{u}_0}V+\mathbf{F}(0), \tag{13}$$

where M is the mass of an ion. Similarly, one may write down the equations of motion for other ions in the lattice.

Now let us assume that the solution of (13) is of the form (9) and let the direction cosines of the vectors $\mathbf{e}$ and $\mathbf{k}$ be denoted by (λ, μ, ν) and (L, M, N) respectively. Then substituting[12] (9) into (13) and making use of (10) and (11), we get

$$\begin{aligned}\lambda\rho\omega^2=\sum_{l,m,n}&[4a^{-2}\{\sin^2\tfrac{1}{2}ka(Ll+Mm+Nn)\}\\&\times\{\delta l(l\lambda+m\mu+n\nu)+\beta\lambda\}]\\&+\frac{\kappa L(\lambda L+\mu M+\nu N)k^2}{1+k^2/(4\pi e\eta)},\end{aligned} \tag{14a}$$

and two similar equations obtained from (14a) by

[10] P. A. M. Dirac, Proc. Cambridge Phil. Soc. **26**, 376 (1930).

[11] See P. Gombas, *Die Statistische Theorie des Atoms und ihre Anwendung* (Springer, Berlin, 1949), p. 78, Eq. (9.5).

[12] As may be seen from symmetry considerations, one would obtain the same equations, namely (14), if one substituted (9) into the equation of motion for an ion whose equilibrium position is at a lattice point not located at the origin of our coordinate system.

TABLE I. Expressions for ρv^2 for long elastic waves propagating along (100), (110), and (111) directions in b.c.c. and f.c.c. crystals.

(LMN)	(λ,μ,ν)	ρv^2 from (14) body-centered cubic	ρv^2 from (14) face-centered cubic	ρv^2 in terms of the elastic constants
(100)	100 (L)[a]	$(8/3)\beta+(8/9)\delta+\kappa$	$4\beta+2\delta+\kappa$	c_{11}
	010 (T)[b]	$(8/3)\beta+(8/9)\delta$	$4\beta+\delta$	c_{44}
	001 (T)	$(8/3)\beta+(8/9)\delta$	$4\beta+\delta$	c_{44}
(110)	110 (L)	$(8/3)\beta+(16/9)\delta+\kappa$	$4\beta+(5/2)\delta+\kappa$	$c_{11}-\frac{1}{2}\epsilon$
	001 (T)	$(8/3)\beta+(8/9)\delta$	$4\beta+\delta$	c_{44}
	1-10 (T)	$(8/3)\beta$	$4\beta+\frac{1}{2}\delta$	$c_{44}+\frac{1}{2}\epsilon$
(111)	111 (L)	$(8/3)\beta+(56/27)\delta+\kappa$	$4\beta+(8/3)\delta+\kappa$	$c_{11}-\frac{2}{3}\epsilon$
	1-10 (T)	$(8/3)\beta+(8/27)\delta$	$4\beta+\frac{1}{3}\delta$	$c_{44}+\frac{1}{3}\epsilon$
	11-2 (T)	$(8/3)\beta+(8/27)\delta$	$4\beta+\frac{1}{3}\delta$	$c_{44}+\frac{1}{3}\epsilon$

[a] L: Longitudinal wave.
[b] T: Transverse wave.

interchanging (l, m, n), (λ, μ, ν), and (L, M, N) as follows:

$$l \rightleftarrows m, \quad \lambda \rightleftarrows \mu, \quad \text{and} \quad L \rightleftarrows M, \tag{14b}$$

and

$$l \rightleftarrows n, \quad \lambda \rightleftarrows \nu, \quad \text{and} \quad L \rightleftarrows N. \tag{14c}$$

Here $\rho(=n_0M)$ is the density of the metal and β, δ and κ are given by

$$\beta=\frac{\rho a^2}{2M}\left(\frac{1}{r}\frac{dW}{dr}\right)_{r=a}, \quad \delta=\frac{\rho a^2}{2M}\left[\frac{d}{dr}\left(\frac{1}{r}\frac{dW}{dr}\right)\right]_{r=a}$$

and

$$\kappa=en_0^2(\eta)^{-1}=en_0^{\frac{2}{3}}\left[\frac{h^2}{4\pi me}\left(\frac{\pi}{3}\right)^{\frac{1}{3}}\times n_0^{-\frac{1}{3}}-0.33en_0^{-\frac{1}{3}}\right]. \tag{15}$$

In deriving (14), use has also been made of the fact that in a cubic crystal corresponding to a nearest neighbor ion at (la, ma, na), there is a nearest neighbor ion at $(-la, -ma, -na)$. The summation in (14), however, is over all the nearest neighbors.

In order that a nontrivial solution ($\mathbf{e}\not\equiv 0$) of (14) exist, the determinant formed from the coefficients of λ, μ, ν must vanish. This gives the secular equation in ω^2 in the form of a 3×3 determinantal equation which will not be written here explicitly. The secular equation is cubic in ω^2 and, therefore, has three solutions (three values of ω^2) for each $\mathbf{k}$. The direction cosines (λ, μ, ν) of the polarization vector $\mathbf{e}$ of the acoustic wave of wave-vector $\mathbf{k}$ and frequency $\omega(\mathbf{k})$ may be determined by substituting these particular values of $\mathbf{k}$ and $\omega^2(\mathbf{k})$ in (14) and then solving the latter for λ, μ, ν.

Identification of the Force Constants β, δ, and κ with the Elastic Constants

For this purpose, we shall solve Eqs. (14) for the velocities $v(=\omega/k)$ of the acoustic waves of infinitely long wavelengths ($k\to 0$). The solution of (14) is particularly simple for waves propagating along any of the three principal directions in a cubic crystal, *viz.*, (100), (110), and (111), since for these three directions of propagation the waves are either wholly longitudinal ($\mathbf{e}\|\mathbf{k}$) or wholly transverse ($\mathbf{e}\perp\mathbf{k}$). In Table I we give expressions for ρv^2 as obtained from (14) for these directions of propagation for both b.c.c. and f.c.c. crystals; corresponding expressions in terms of the elastic constants are given in the last column of Table I ($\epsilon=c_{11}-c_{12}-2c_{44}$).

For each type of lattice structure, we have from Table I nine relations between the three force constants β, δ, and κ and the three elastic constants. However, they are not all independent of each other and, as may be readily verified, they are identically satisfied if we put

$$\left.\begin{aligned}\beta&=\tfrac{3}{8}(c_{44}+\tfrac{1}{2}\epsilon)\\ \delta&=-(9/16)\epsilon\\ \kappa&=c_{11}-c_{44}\end{aligned}\right\}\text{for b.c.c.,} \tag{16a}$$

and

$$\left.\begin{aligned}\beta&=\tfrac{1}{4}(c_{44}+\epsilon)\\ \delta&=-\epsilon\\ \kappa&=c_{11}-c_{44}+\epsilon\end{aligned}\right\}\text{for f.c.c.} \tag{16b}$$

Moreover, by substituting (16a) or (16b) in (14) and remembering that the body-centered and face-centered cubic lattices have eight and twelve nearest neighbors respectively, one may verify that for the limiting case of infinitely long waves ($k\to 0$), Eqs. (14) reduce to the Christoffel equations of the elasticity theory.[13]

From (16), we also have

$$\left.\begin{aligned}c_{11}-c_{12}&=(16/3)\beta\\ c_{44}&=(8/3)\beta+(8/9)\delta\end{aligned}\right\}\text{for b.c.c.,} \tag{17a}$$

and

$$\left.\begin{aligned}c_{11}-c_{12}&=8\beta+\delta\\ c_{44}&=4\beta+\delta\end{aligned}\right\}\text{for f.c.c.} \tag{17b}$$

Thus c_{44} and $c_{11}-c_{12}$ are independent of κ, and are completely determined by the interaction potential[14] $W(r)$, a result which is in agreement with (1).

By substituting (16a) or (16b) in (14) and inserting the observed elastic constants for any particular metal, one may solve the latter for ω^2 for any $\mathbf{k}$. Further, if one defines a frequency distribution function $G(\omega)$ such that $G(\omega)d\omega$ denotes the number of frequencies lying between ω and $\omega+d\omega$, one may calculate $G(\omega)$ by solving (14) for sufficiently large number of points in k-space. The distribution function $G(\omega)$ for sodium, and hence its specific heat, will be calculated in Sec. III. We shall, however, first discuss some of the distinctive features of Eqs. (14) and the limitations inherent in the model on whose basis they have been derived.

Discussion

First, it will be noticed from (16) that in the present treatment one linear combination of the elastic constants—$\kappa=c_{11}-c_{44}$ and $\kappa=c_{11}-c_{44}+\epsilon$ for body- and face-centered cubic lattices respectively—depends es-

[13] A. E. H. Love, *Mathematical Theory of Elasticity* (Dover Publications, New York, 1944), fourth edition, p. 299.

[14] It may be verified that expressions (17) for $c_{11}-c_{12}$ and c_{44} are the same as those obtained by considering changes in the potential energy $\Sigma W(\mathbf{r})$ of a crystal due to appropriate *static* homogeneous strains; this provides a check on our calculations.

sentially only on the number of free electrons per atom in the metal. Taking this number to be 1 per atom, the calculated values of κ for Na (b.c.c.) and Ag (f.c.c.) are 0.3×10^{11} and 1.7×10^{11} dynes/cm^2 respectively as compared with the corresponding experimental values 0.39×10^{11} and 2×10^{11} dynes/cm^2. The agreement between the two values for other monovalent metals is not so satisfactory, the worst examples being Li and Cu for which κ_{cal} is 0.99×10^{11} and 3.5×10^{11} dynes/cm^2 as compared with the experimental values 0.23×10^{11} and 0.2×10^{11} dynes/cm^2 respectively. Of course, in view of the crudeness of the model adopted here, and, in particular, since the use of the Sommerfeld model for a metal implies that the effect of the energy E_0 (E_0 depends on the atomic volume only) of the ground state of the valence electron on the ionic motion, and hence on the bulk modulus $\frac{1}{3}(c_{11}+2c_{12})$ of the crystal, is negligible, one can hardly expect more than a qualitative agreement between the experimental and calculated values of κ. Indeed, if in the expression (15) for κ, the second term, which arises from the exchange interaction between the electrons, is neglected, κ becomes

$$\kappa=(h^2/4\pi m)(\pi/3)^{\frac{1}{3}}\times n_0^{5/3}; \tag{18}$$

the right-hand side of (18) may be immediately recognized as the bulk modulus K of the electron gas if one remembers that $K=-V(\partial p/\partial V)$ and that the pressure p of an electron gas enclosed in a volume V is given by the expression[15] ($n=n_0V$)

$$p=(3h^2/20\pi m)(\pi/3)^{\frac{1}{3}}(n/V)^{5/3}. \tag{19}$$

Secondly, since the metal is treated here partly as a continuous system, the secular equation (14), and hence the frequencies ω obtained from it, will generally not be periodic functions of $\mathbf{k}$. We have, therefore, to restrict $\mathbf{k}$ values in some suitable way so that the total number of normal modes of vibration exactly equal the number of degrees of freedom, *viz.*, three per atom. Now, had the periodicity of the ionic lattice not been ignored, (14) and ω would have been distinct only for $\mathbf{k}$ values lying inside the first Brillouin zone in k-space, and the solutions corresponding to a $\mathbf{k}$ lying outside this zone would have given no new normal modes of vibration. Since the first Brillouin zone has just one $\mathbf{k}$-point per atom, one obtains the correct total number of normal modes of vibration by restricting the k values to the first Brillouin zone.[16] Therefore, it would seem reasonable to restrict the k values to the first Brillouin zone while calculating the frequency spectrum from (14) also. It may be mentioned that since the forces arising from the energies E do not contribute to the motion of the *wholly* transverse waves, the frequencies corresponding to these waves will be periodic functions of $\mathbf{k}$ [see Eqs. (28)].

Lastly, the application of the Thomas-Fermi method to calculate changes in φ due to *time-dependent* ionic displacements contains the implicit assumption that the electrons follow the motion of the ions "adiabatically" in the sense of the Thomas-Fermi method—namely, that the electrons occupy the lowest energy state of the instantaneous lattice potential. This seems to be reasonable for low frequencies. For higher frequencies, the validity of this assumption may be doubted. However, it does not appear easy to make a decision on this point from a comparison of the theoretical and observed specific heats since, first, the high frequency end of the spectrum affects the specific heats at comparatively high temperatures only and that not very sensitively, and secondly, the accuracy of the theoretical values is limited by the approximations involved in the calculation of frequency spectrum and by the accuracy of the available observed elastic constants which have to be inserted in the secular equation (14).

Although the foregoing discussion makes it clear that a more refined treatment than given here of the problem of determining the frequencies of the normal modes of vibration of a metal would be desirable, the present approach is likely to be an improvement over the customary two force-constant model. This is because, unlike the latter, it is able to incorporate all the three independent elastic constants of a cubic metal in the secular equation. The uncertainties mainly arise at the high frequency end of the spectrum, and these do not affect the results on specific heats very sensitively. Thus one may expect that, provided one uses the actually observed elastic constants in the secular equation (14), the calculations of frequency spectrum, and hence of specific heats, based on (14) would be an improvement over those based on a two force-constant model. This, of course, would be all the more so for the low-frequency end of the spectrum and hence for the specific heats at low temperatures.

II. FREQUENCY SPECTRUM AND SPECIFIC HEAT OF SODIUM

In this section, the frequency spectrum of sodium, and hence its specific heat C_v at constant volume, will be calculated from the solutions of the secular equation (14).

[15] See, for example, reference 11, p. 6.

It should be mentioned here that recently Jules de Lunay [J. Chem. Phys. **29**, 1975 (1953)] has made an attempt to make a three force-constant model for a metal. Lunay superimposes the free electron gas on the Born-Kármán model, but appears to assume that the ionic lattice is in equilibrium under the central forces alone. He thus obtains the result that the deviation $c_{12}-c_{44}$ from Cauchy's relation is just equal to the bulk modulus of the free electron gas. No calculations of frequency spectrum have been made on this model. The author is indebted to Dr. E. Bauer for drawing his attention to this paper.

[16] For a periodic lattice, these results are obtained by the use of Born's cyclic boundary conditions. It should be mentioned that in general there exist several equivalent portions of reciprocal space to which $\mathbf{k}$ vectors may be confined. Of these, the Brillouin zones have the maximum possible geometrical symmetry consistent with that of the lattice. We choose one of these zones (namely, the *first* for obvious reasons) because the approximate method of Houston, by which the frequency spectrum is calculated in Sec. III, is the more reliable the more symmetrical is the k-space to which $\mathbf{k}$ vectors are confined.

Houston's Method of Calculating Frequency Spectrum

For cubic crystals, Houston's[17] method of calculating the frequency spectrum is, by far, the most convenient one. There exist for a cubic crystal certain directions (θ_n,ϕ_n) in k-space along which the secular equation in ω^2 can be solved analytically. Corresponding to each $\mathbf{k}$ there are, of course, three frequencies corresponding to the three branches. In the following, these will be distinguished by the subscript i $(i=1, 2, 3)$. The number $g(\omega)d\omega d\Omega$ of normal modes (frequencies) which lie between ω and $\omega+d\omega$ and whose wave vectors $\mathbf{k}$ lie in a solid angle $d\Omega$ around $\mathbf{k}(\theta,\phi)$ may be written as

$$g(\omega,\theta,\phi)d\omega d\Omega=\sum_i g_i(\omega,\theta,\phi)d\omega d\Omega$$
$$=\frac{V}{(2\pi)^3}\sum_i k^2\frac{dk}{d\omega_i(\theta,\varphi)}d\omega d\Omega, \quad (20)$$

where V is the volume of the crystal under consideration. One now expands $g_i(\omega,\theta,\phi)$ in terms of those Kubic Harmonics[18] $K_m(\theta,\phi)$ which are invariant under the operations of the cubic symmetry group; thus

$$g_i(\omega,\theta,\phi)=\sum_m a_{i,m}K_m(\theta,\phi), \quad (K_0=1). \quad (21)$$

The coefficients $a_{i,m}$ in (21) may be determined from the solutions of the secular equation along the various directions (θ_n,ϕ_n) in k-space. The total number of frequencies $G(\omega)d\omega$ lying between ω and $\omega+d\omega$ is then obtained from (21) by integrating it over all solid angles and summing over the three branches. Remembering that the various K_m are orthogonal to each other, we have

$$G(\omega)d\omega\equiv\sum_i G_i(\omega)d\omega=\sum_i d\omega\int g_i(\omega,\theta,\phi)d\Omega$$
$$=4\pi d\omega\sum_i a_{i,0}. \quad (22)$$

This method naturally gives the better approximation to $G(\omega)$ the more numerous are the directions (θ_n,ϕ_n) along which the secular equation is solved. In practice, one solves the secular equation along the three principal directions of a cubic crystal, *viz.*, (100), (110), and (111) [the corresponding direction cosines being (100), $2^{-\frac{1}{2}}(110)$, and $3^{-\frac{1}{2}}(111)$]; the quantities referring to these three directions will be distinguished by subscripts A, B, and C respectively. Then one has to retain only the first three terms in the expansion (21). Apart from normalization factors, the first three appropriate Kubic Harmonics, expressed in Cartesian coordinates, are $(x^2+y^2+z^2=1)$

$$K_0=1, \quad K_1=(x^4+y^4+z^4-\tfrac{3}{5}),$$
$$K_2=x^2y^2z^2+(1/22)K_1-(1/105). \quad (23)$$

Making use of (20) to (23), one obtains for $G(\omega)$ the expression[19]

$$G(\omega)=\frac{V}{(2\pi)^3}\frac{4\pi}{35}\sum_i\Big[10k^2\Big(\frac{dk}{d\omega}\Big)_{i,A}$$
$$+16k^2\Big(\frac{dk}{d\omega}\Big)_{i,B}+9k^2\Big(\frac{dk}{d\omega}\Big)_{i,C}\Big]. \quad (24)$$

In general, the right-hand side of (24) has to be multiplied by a normalization factor so that one obtains the correct number of normal modes, *viz.*,

$$\int G_i(\omega)d\omega=N_0V \quad (i=1, 2 \text{ and } 3), \quad (25)$$

where N_0 is the number of atoms per unit volume. If we denote the maximum value of k along the three directions A, B, and C by $k_A(m)$, $k_B(m)$ and $k_C(m)$ respectively, the normalization factor F is determined by the equation

$$\frac{F}{(2\pi)^3}V\cdot\frac{4\pi}{35}\tfrac{1}{3}[10k_A^3(m)+16k_B^3(m)+9k_C^3(m)]$$
$$=N_0V. \quad (26)$$

As already mentioned in Sec. II, the $\mathbf{k}$ vectors are to be restricted to the first Brillouin zone. For a body-centered cubic lattice—this is the case we are interested in here since sodium has body-centered cubic structure—the first Brillouin zone is a dodecahedron bounded by (110) planes. The distances of these planes from the origin of k-space are given in reference 4, p. 156, and we may determine $k_A(m)$, $k_B(m)$, and $k_C(m)$ by finding the points of intersection of the corresponding $\mathbf{k}$ vectors with (110) plane. One finds[20]

$$k_A(m)=\sqrt{3}(\pi/a), \quad k_B(m)=(\sqrt{\tfrac{3}{2}})(\pi/a),$$
$$k_C(m)=\tfrac{3}{2}(\pi/a). \quad (27)$$

From (26) and (27), one obtains for the normalization factor F:

$$F=0.778.$$

Before we use (24) to calculate the frequency spectrum of sodium, we should mention that the frequency spectrum obtained by Houston's method outlined above contains, in general, a certain number of singularities; the area under them, however, is finite. These singularities are peculiar to Houston's method and arise from the fact that the frequency spectrum is approximated from the solutions of the secular equation along three directions only. Van Hove[21] has investigated

[17] W. V. Houston, Revs. Modern Phys. **20**, 161 (1948). An account of various methods of calculating the frequency spectrum may be found in reference 1, pp. 70–84.

[18] Kubic Harmonics were introduced by F. C. Von der Lage and H. A. Bethe, Phys. Rev. **71**, 612 (1947).

[19] This is essentially Houston's expression (11) after allowing for a misprint in his equation preceding (11), where the coefficient of $f_1(q)$ should read $-2/(4\pi)^{\frac{1}{2}}$ instead of $-10/(36\pi)^{\frac{1}{2}}$.

[20] Note that in reference 4, a is the lattice constant while here a is distance between the nearest neighbors; moreover, k as defined here is 2π times the k used in Chapter V of reference 4.

[21] L. Van Hove, Phys. Rev. **89**, 1189 (1953).

the nature of the singularities of $G(\omega)$ in one, two, and three dimension. He finds that in three dimensions $G(\omega)$ can only have finite discontinuities. Nakamura,[22] on the basis of an analysis of two dimensions, concludes that Houston's method reproduces fairly accurately the shape of the frequency spectrum up to the first maximum; beyond this, the actual $G(\omega)$ is flatter than that obtained by the Houston method. Since the numerical work involved in systematically improving Houston's approximation is rather large and since the slight flattening of the distribution function beyond the first maximum is likely to affect, if at all, the calculated specific heats in the temperature region $T\sim\Theta$ (effective Debye temperature) only, we shall use here Houston's method. The effect of the flattening of $G(\omega)$ will be discussed qualitatively at the end of this section. For the purposes of numerical integrations, however, the singularities will be replaced by finite peaks of equal area.

Frequency Spectrum

The first step towards obtaining the distribution function $G(\omega)$ is to solve equations (14) for ω^2 along the three directions A, B, and C. The solution is elementary; making use of (16a), we find

Direction A

$(\lambda,\mu,\nu)=(100)$:

$$\omega^2=\frac{12}{a^2\rho}c_{44}\sin^2\left(\frac{ka}{2\sqrt{3}}\right)+\frac{(c_{11}-c_{44})k^2}{\rho[1+k^2/(4\pi e\eta)]}, \quad (28a)$$

$(\lambda,\mu,\nu)=(010)$ or (001):

$$\omega^2=\frac{12}{a^2\rho}c_{44}\sin^2\left(\frac{ka}{2\sqrt{3}}\right). \quad (28b)$$

Direction B

$(\lambda,\mu,\nu)=(1,1,0)$:

$$\omega^2=\frac{1}{a^2\rho}(6c_{44}-3\epsilon)\sin^2\left(\frac{ka}{\sqrt{6}}\right)+\frac{(c_{11}-c_{44})k^2}{\rho[1+k^2/(4\pi e\eta)]}, \quad (28c)$$

$(\lambda,\mu,\nu)=(1,-1,0)$:

$$\omega^2=\frac{1}{a^2\rho}(6c_{44}+3\epsilon)\sin^2\left(\frac{ka}{\sqrt{6}}\right), \quad (28d)$$

$(\lambda,\mu,\nu)=(001)$:

$$\omega^2=\frac{6}{a^2\rho}c_{44}\sin^2\left(\frac{ka}{\sqrt{6}}\right). \quad (28e)$$

[22] T. Nakamura, Progr. Theoret. Phys. (Japan) 5, 213 (1950).

Direction C

$(\lambda,\mu,\nu)=(1,1,1)$:

$$\omega^2=\frac{1}{a^2\rho}\left[3(c_{44}-\epsilon)\sin^2\left(\frac{ka}{2}\right)+3(3c_{44}+\epsilon)\sin^2\left(\frac{ka}{6}\right)\right]+\frac{(c_{11}-c_{44})k^2}{\rho[1+k^2/(4\pi e\eta)]}, \quad (28f)$$

$(\lambda,\mu,\nu)=(1,-1,0)$ or $(1,1,-2)$:

$$\omega^2=\frac{1}{a^2\rho}\left[3(c_{44}+\tfrac{1}{2}\epsilon)\sin^2\left(\frac{ka}{2}\right)+3(3c_{44}-\tfrac{1}{2}\epsilon)\sin^2\left(\frac{ka}{6}\right)\right]. \quad (28g)$$

It will be seen from (28) that, just as in the case of infinitely long waves (see Sec. II), the elastic waves of any wave number k propagating along one of these three directions are either wholly longitudinal or wholly transverse. In other directions, there is a slight mixing up of the longitudinal and transverse vibrations. Hence the frequency distribution obtained from (28a), (28c), and (28f) will be predominantly due to the longitudinal waves, while that obtained from the remaining equations in (28) will be chiefly due to the transverse waves. We denote these two distribution functions by $G_l(\omega)$ and $G_t(\omega)$ respectively. Obviously

$$G(\omega)=G_l(\omega)+G_t(\omega) \quad \text{and}$$

$$\int G_l(\omega)d\omega=\frac{1}{2}\int G_t(\omega)d\omega=N_0V.$$

The frequency distribution functions $G_l(\omega)$, $G_t(\omega)$ and $G(\omega)$ are now easily obtained with the help of (24), (27), and (28). The elastic constants for sodium used in this calculation were the same as employed by Bauer.[7] These, as well as other constants needed in the calculation, are listed below:

$$\begin{gathered}c_{44}=0.580\times10^{11},\quad c_{11}-c_{44}=0.390\times10^{11},\\ \epsilon=-1.02\times10^{11}, \text{ all of them in dynes/cm}^2;\\ n_0=N_0=2.56\times10^{22}/\text{cc},\quad a=(3\sqrt{3}/4n_0)^{\frac{1}{3}},\\ \rho=0.97 \text{ g/cc},\end{gathered} \quad (29)$$

$$(4\pi e\eta)^{-1}=\frac{1}{4\pi e}\cdot\frac{M}{\rho}\cdot\frac{(c_{11}-c_{44})}{en_0}=0.015\times a^2 \quad [\text{see Eq. (16)}].$$

The distribution functions $G_l(\omega)$, $G_t(\omega)$, and $G(\omega)$ are plotted against ω in Fig. 1.

Specific Heat of Sodium

From the distribution function $G(\omega)$, the specific heat C_v at constant volume may be obtained in the

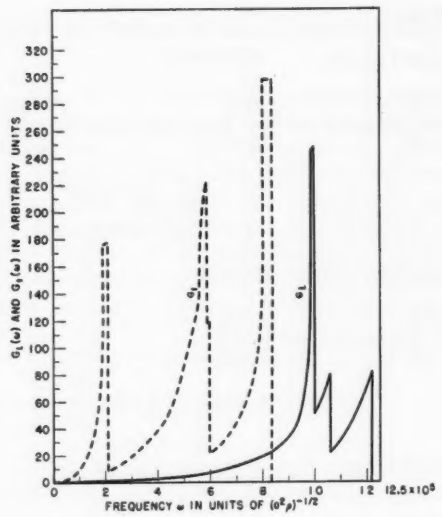


FIG. 1. The frequency spectrum of sodium. The distribution functions $G_l(\omega)$ and $G_t(\omega)$, predominantly due to longitudinal and transverse waves respectively, are given in Fig. 1(a), while $G(\omega)=G_l(\omega)+G_t(\omega)$ is given in Fig. 1(b). (For convenience, the singularities have been replaced by finite peaks of equal area.) The normalized distribution functions may be obtained by multiplying the ordinates in Fig. 1 by $3.9\times10^{-8}(a^2\rho)^{\frac{1}{2}}N_0V$.

usual manner by numerical integration. The results of the entire numerical work of the present paper are estimated to be correct to within about 3 percent. The calculated and experimental specific heats are plotted against the temperature T in Fig. 2. The experimental values of C_v above 20°K are due to Simon and Zeidler,[23] while those below this temperature are due to Pickard and Simon.[24]

A more usual way to study the C_v-T behavior is to insert C_v at a given T in the Debye formula for specific heats, calculate the characteristic temperature Θ and plot Θ against T. If $G(\omega)$ is strictly proportional to ω^2, Θ is independent of temperature; otherwise Θ depends on T. In Fig. 3, we give $\Theta-T$ curves for (*a*) Θ derived from the C_v calculated in the present paper, (*b*) Θ derived from the observed specific heats. As mentioned in the introduction, Bauer,[7] using a two force-constant model, has calculated the specific heat of sodium. For comparison, we give in Fig. 3, curve *c*, Θ's obtained by him on Houston's approximation and finally curves (*d*) and (*e*) which give Θ's obtained by Bauer by flattening in two different ways[25] the frequency distribution function $G(\omega)$ calculated by Houston's method. We should mention that for temperatures above about 120°K, the theoretical Θ-values [presented in the dotted part of the curve (*a*)] are to be regarded as giving merely the general shape of the curve since, at these temperatures, a small error in C_v causes a large error in the corresponding Θ value.

It will be seen from Fig. 3 that the agreement be-

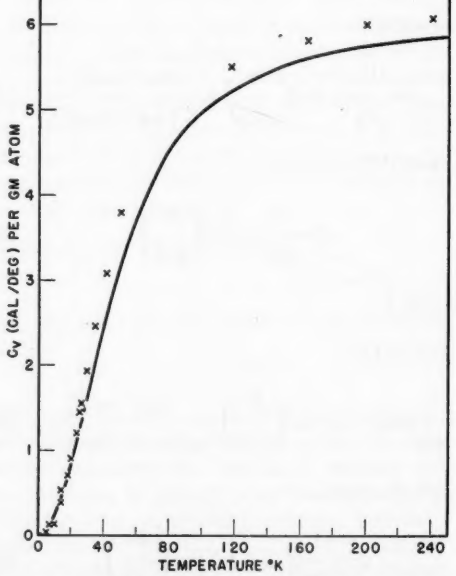


FIG. 2. The calculated and observed lattice specific heats C_v of sodium, as a function of temperature. The experimental points are shown by X. (The electronic specific heat, which was subtracted at low temperatures from the observed C_v, is assumed to be given by the formula to be found in any textbook on the subject.)

[23] F. E. Simon and W. Zeidler, Z. physik. Chem. **B123**, 383 (1926).

[24] G. L. Pickard and F. E. Simon, Proc. Phys. Soc. (London) **A61**, 1 (1948).

[25] See reference 7, p. 63.

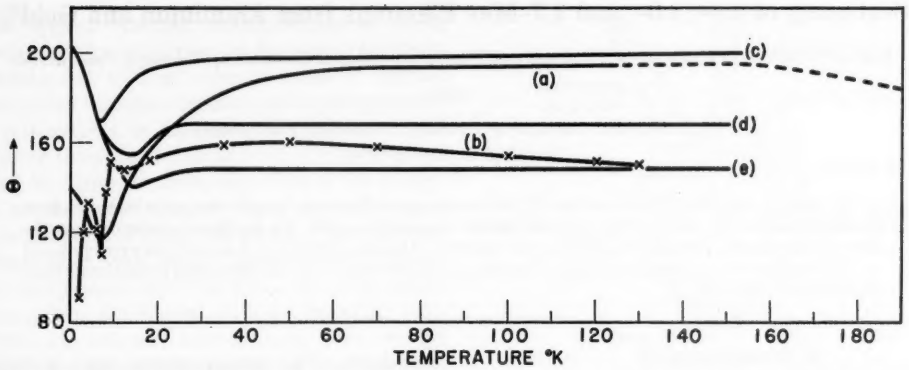


FIG. 3. The effective Debye temperature Θ for sodium, as a function of temperature. Curve (*a*): present calculations; curve (*b*): experimental. Curves (*c*), (*d*) and (*e*) reproduce the theoretical results of Bauer[7]; the explanation of these curves is given in the text.

tween the calculated and observed values of Θ [curves (*a*) and (*b*)] is satisfactory[26] below about 30°K; the discrepancy below about 4°K is presumably due to the uncertainty in the contribution of the electronic specific heat. This specific heat, according to recent investigations by Buckingham,[27] may not be proportional to T as is usually assumed. Above 30°K, the theoretical values are about 15 percent higher than the experimental values. However, it will be noticed that the curve (*a*) lies everywhere much closer to the experimental curve than the curve (*c*) which was obtained by Bauer[7] using Houston's method, i.e., by the same method of calculating the frequency spectrum as employed here. The discrepancy at higher temperatures between the Θ-values calculated here and the experimental values may be due to one or more of the following causes:

(1) The actual frequency distribution beyond the first maximum is flatter than that obtained on Houston's approximation. This will lower the theoretical values at higher temperatures. For example, the curves (*d*) and (*e*) of Bauer lie lower than his curve (*c*).

(2) The assumptions on which the secular equation (14) is derived make the high frequency end of the spectrum somewhat uncertain (see discussion at the end of Sec. II).

(3) Temperature variation of the elastic constants which is neglected in the present calculation. (The values used in the calculation refer to 0°K).

(4) Finally, it will be seen from Fig. 2 that the observed C_v rises above the classical value 6 cal per degree per g atom at about 200°K, while the C_v calculated from the frequency spectrum of normal modes of vibration can never rise above this value. This implies that part of the contribution to the experimental specific heat at this temperature, and therefore possibly at lower temperatures too, comes from sources other than elastic vibrations, e.g., from anharmonicity. If this extraneous specific heat is allowed for, the experimental points in Fig. 3 would tend to lie, at least above about 100°K, a little higher than at present.

Conclusion

From the foregoing discussion we may conclude that, on the whole, the agreement between the experimental values of C_v and those calculated here is satisfactory. In particular, the close agreement between the two sets of values below about 30°K indicates the necessity of using, as has been done in the present paper, some three force-constant model for a metal like sodium which is neither elastically isotropic nor satisfies the Cauchy relation $c_{12}=c_{44}$.

ACKNOWLEDGMENTS

It is a great pleasure to thank Dr. T. Y. Wu for his interest in this work and to Drs. R. B. Dingle and G. K. Horton for helpful discussions. The author is also indebted to the I.C.I. Committee of the University of Edinburgh for granting him leave of absence and to the National Research Council, Canada, for an award of a Fellowship.

[26] We should mention that although both curves (*b*) and (*a*) show a minimum at about 7°K, the minimum in the experimental curve is deeper than that in the theoretical curve; this deep minimum in the former is due to a slight hump at about 7°K in the experimental C_v-T curve of reference 24 and does not appear to have been confirmed by the recent work of Hill, Smith, and Parkinson (unpublished); see M. J. Buckingham and M. R. Schafroth, Proc. Phys. Soc. (London) **A67**, 828 (1954), footnote on p. 829. The C_v-T curve obtained from the present calculations has no hump.

[27] M. J. Buckingham, Nature **168**, 281 (1951); see also the reference given in the previous footnote.

Scattering of 0.6-, 1.0-, and 1.7-Mev Electrons from Aluminum and Gold*†

ROBERT T. BAYARD,‡ *University of Pittsburgh and Westinghouse Electric Corporation, Pittsburgh, Pennsylvania*

AND

J. L. YNTEMA, *University of Pittsburgh, Pittsburgh, Pennsylvania*

(Received October 4, 1954)

The relative differential cross sections for the scattering of electrons by gold and aluminum has been measured at 0.6, 1.0, and 1.7 Mev over the angular range 30° to 150°. No significant deviation from the theoretically expected angular distribution was observed. Absolute differential cross sections were measured at 60° for aluminum at the same energies. The data are in agreement with the theoretical predictions within the experimental error.

I. INTRODUCTION

THE differential cross sections for the scattering of relativistic electrons by nuclei has been evaluated by Mott.[1] The earlier experimental results have been summarized by Mott and Massey.[2] These early results present a rather confused picture. More recent measurements by Van de Graaff and co-workers,[3] however, gave good agreement with the theoretical predictions for the absolute differential cross section at energies from 1.27 to 2.27 Mev and for a variety of elements. Their measurements extended only over an angular range of 20° to 50°. Measurements extending from 30° to 150° on 0.245-Mev electrons were reported in 1952 by Kinzinger and Bothe,[4] and from 60° to 120° on 2.2-Mev electrons by Paul and Reich.[5] Subsequent measurements at 120° were reported by Kinzinger[6] between 0.15 and 0.4 Mev. In these last experiments good agreement with theory at all angles was found for aluminum, but rather sizable deviations at large angles were found for the heavier elements gold and platinum. Since the theory is expected to describe the situation reasonably well at these energies, and since no cause could readily be found to explain the experimental deviations, it seemed desirable to make another measurement of relative electron scattering.

In the experiment of Kinzinger and Bothe, the relative scattering cross section ratios in aluminum at 30°/50°, 50°/70°—130°/150° were first measured, and then the cross sections for gold, silver and nickel at various angles were compared with those obtained for aluminum. The measurements were normalized by measuring the target foil thicknesses and keeping the electron beam current constant. It was assumed that at all angles the backscattering from the interior walls of the scattering chamber into the detectors would be negligible. Paul and Reich used a somewhat similar experimental arrangement.

It seemed desirable in the present experiment to avoid errors in the relative scattering measurements due to uncertainties in the foil thickness determinations and in the constancy of the beam current. Therefore relative scattering measurements were made directly in gold, and independently, in aluminum. The measurements in aluminum were made principally to check the operation of the apparatus. The scattering intensities at 30°, 60°, 120°, and 150° were each compared with that at 90° in order to avoid the cumulative errors associated with a sequential comparison of scattering ratios 30°/60°, 60°/90°, etc. The energies for which measurements were made were 0.6, 1.0, and 1.7 Mev. It was found that the assumption made by Kinzinger and Bothe, that backscattering of electrons into the detectors was unimportant, could not be made for the present experimental arrangement. Multiple-scattering corrections could not be neglected entirely at small angles, but they were made small through the use of thin target films.

In order to provide another measurement of the absolute scattering cross section, as well as to check the operation of the experimental apparatus, a determination of absolute scattering was made in aluminum at 60° for electrons of the three energies used in the relative scattering measurements.

II. APPARATUS

A. Electron Source

The source of electrons used was a 2-Mev Van de Graaff generator. The beam was brought to a focus at the scattering foil location by means of a lens-type focusing coil. The diameter of the major part of the focused beam varied from about 5 mm for 0.6 Mev to 3 mm for 1.0 and 1.7 Mev. These diameters were determined by permitting the beam to pass through a piece of clear Scotch tape, which turns white in the

* This work was supported in part by a joint program of the Office of Naval Research and the U. S. Atomic Energy Commission.

† Submitted to the University of Pittsburgh by R. T. Bayard as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

‡ Present address: Westinghouse Atomic Power Division, Pittsburgh, Pennsylvania.

[1] N. F. Mott, Proc. Roy. Soc. (London) **A124**, 425 (1929).

[2] N. F. Mott and H. S. W. Massey, *The Theory of Atomic Collisions* (Oxford University Press, London, 1949), second edition.

[3] Van de Graaff, Buechner, and Feshbach, Phys. Rev. **69**, 452 (1946); Buechner, Van de Graaff, Sperduto, Burrill, and Feshbach, Phys. Rev. **72**, 678 (1947).

[4] E. Kinzinger and W. Bothe, Z. Naturforsch. **7a**, 390 (1952).

[5] W. Paul and H. Reich, Z. Physik **131**, 326 (1952).

[6] E. Kinzinger, Z. Naturforsch. **8a**, 312 (1953).

irradiated region. Another determination of beam spread was made by measuring the fraction of the beam which would fail to pass through holes of various diameters drilled through aluminum plates and centrally located with respect to the beam. It was determined that no more than about 0.05 percent of the beam fell outside a $\frac{3}{8}$-in. diameter circle, regardless of the energy. Because of the large distance from the collimating magnet to the point of focus, the paths of the electrons in the beam were very nearly parallel, the maximum deviation being about 0.6° (for 0.6 Mev). The energy of the beam was determined with a generating voltmeter which was also used to control the beam energy. Calibration was achieved by measuring the beryllium and deuterium gamma-neutron thresholds of (1.666 ±0.002) Mev and (2.227±0.003) Mev.[7] The beam energy would fluctuate in some cases as much as ±1 percent, although it was usually more stable. The soft component of the beam, as determined by the current measured with the electron emitting filament off, was not measurable with a meter which could detect currents of the order of 10^{-10} amperes. The beam currents used varied from a few times 10^{-9} amperes to 10^{-6} amperes.

B. The Scattering Chamber and Accessories

A sectional view of the scattering chamber is shown in Fig. 1. Some of the details in this figure are not strictly to scale. The chamber and all parts exposed to the direct or scattered beam are aluminum. The main body of the chamber is an aluminum ring $1\frac{3}{4}$ in. thick and 11 in. i.d. with machined ports and faces located every 30°. This central ring has two aluminum coverplates which are sealed by means of O-ring gaskets. The ports used for entry and exit of electrons during measurements have appropriate fittings attached and sealed by O-ring gaskets, while those ports not in use have aluminum covers.

The purpose of the collimators to which the detectors are attached is to reduce the number of electrons which can reach the detectors without having been scattered directly from the foil. The diaphragms of the collimators are $\frac{1}{8}$-in.-thick aluminum and the holes are $\frac{3}{8}$ in. in diameter. The face and shoulder which mate with the face and port of the scattering chamber are machined with respect to the diaphragm holes so that when the collimator is bolted in place, the centerline of the diaphragms passes through the center of the chamber. The diaphragm farthest from the scattering foil, i.e., that adjacent to the detector crystal, constitutes the aperture which defines the scattering solid angle. This beam defining aperture is covered by a vacuum tight window of 0.001-in. aluminum through which the scattered electrons pass.

As long as the region of the foil in which scattering

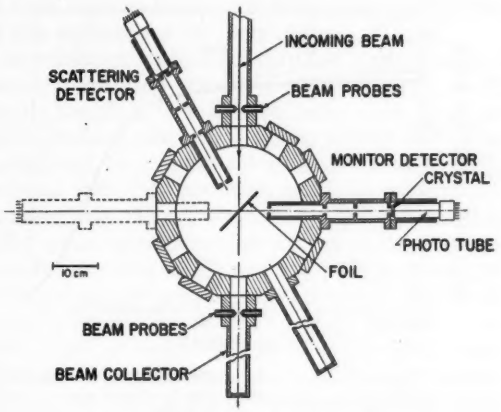


FIG. 1. The scattering chamber.

occurs is smaller than the diameter of the collimating diaphragm holes, these diaphragms will not interfere with the electrons scattered in the solid angle of the defining aperture. However, they do stop a large number of electrons which are backscattered from various parts of the chamber walls, many of which have very nearly the full energy of the initial beam. The diaphragms cannot, of course, stop electrons which are scattered from points directly opposite a collimator, pass undeflected through the foil, and into the detector. The effect was particularly important at scattering angles of 120° and 150° because of the relatively large numbers of electrons originally scattered at 60° and 30° respectively and backscattered into the detectors. This was even more serious with short (six-in.) collimators which were used in the early phases of the experiment. These short collimators were essentially the same as those shown in Fig. 1 except that they did not extend into the scattering chamber and had only three diaphragms. This backscattering was made negligible by (a) making the collimator longer ($10\frac{1}{2}$ in.), as shown in Fig. 1, so that the area of the chamber wall which could be seen by the detector was reduced, and (b) by locating a long aluminum tube opposite the detector (shown in Fig. 1 at the 30° position). Thus the electrons originally scattered at the small angle do not encounter any material from which to backscatter until they are a considerable distance from the chamber (about four feet) and the probability for backscattering into the detector is reduced considerably. The effect of these changes in geometry on the pulse height distribution can be seen in Fig. 2. Composites of several curves are given for the original geometry at 150° (upper curve), for the improved geometry at 150° (middle curve) and for the improved geometry at angles from 30° to 120°. For the range of pulse height shown, the curves for 30° to 120° are indistinguishable. For pulse heights in the neighborhood of that representing the beam energy, the 150° curve also coincided with the others. For lower pulse heights

[7] R. C. Mobley and R. A. Laubenstein, Phys. Rev. 80, 309 (1950).

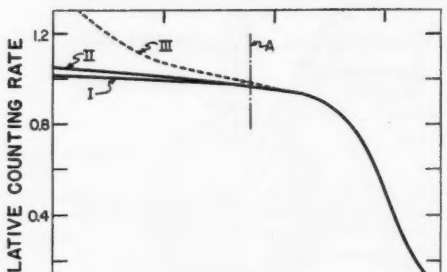


FIG. 2. Typical integral pulse height distributions obtained in relative scattering measurements. Curve I—Composite of characteristics from 30° to 120°, Curve II—150° with improved collimator and chamber geometry, Curve III—150° with original geometry.

there is a slight difference for the improved geometry case, and a considerable difference for the original geometry.

The foils were mounted on an oval aluminum frame with inside dimensions of 1.25 in.×2 in. The frames were mounted on a shaft which passed through one of the scattering chamber side plates and which could be rotated from outside the chamber. The side plate was accurately positioned so that the axis of the shaft coincided with the center of the chamber. A special side plate with provision for positioning either of two foil frames in the center of the chamber was used to determine the background from x-rays, etc. In these measurements, one of the frames had a foil and the other was blank.

The aluminum foils used were commercially prepared and had a thickness of about 0.2 mg/cm². The measurements on aluminum at 1.0 Mev and those for 30° and 60° at 0.6 Mev were made with a foil consisting of a foil support of ~60 μg/cm² Parlodion supporting a $\frac{3}{8}$-in. diameter spot of the 0.2-mg/cm² aluminum in the center. The original purpose of providing the scatterer only in the center was to eliminate any scattering outside the $\frac{3}{8}$-in. diameter spot which could be seen by the detector. Since subsequent measurements showed that a negligible amount of the beam fell outside a $\frac{3}{8}$-in. circle, the remaining measurements of aluminum scattering at 0.6 Mev and those at 1.7 Mev were made with an unbacked foil across the entire frame. The gold scatterer consisted of a spot of gold $\frac{3}{8}$ in. in diameter evaporated on the center of a Parlodion foil support. The gold was thin enough to transmit visible light. The thickness determined from the absolute scattering at 60° was ~16 μg/cm², while that estimated by assuming isotropic evaporation and 100 percent sticking of the gold atoms on the Parlodion was ~20 μg/cm², or about 25 percent higher. It is interesting in this connection to note that Chase and Cox[8] obtained an estimate of the thickness of an evaporated aluminum foil by assuming isotropic evaporation which was 19 percent higher than the value they obtained by microchemical analysis of the same foil.

The detectors were scintillation counters each consisting of a $\frac{1}{4}$-in.-thick×$1\frac{1}{2}$-in.-diameter anthracene crystal and a 2-in.-diameter type 6292 photomultiplier. The pulses from the photomultipliers were fed into linear amplifiers and pulse height discriminators and then into scale-of-1000 scalers. The scalers were switched on and off simultaneously so that the counts in the two channels were recorded over the same time interval. The pulse height resolution, as determined with a single-channel pulse height analyzer, was of the order of 10 percent for all detectors.

The alignment of the beam was accomplished by measuring the beam position with probes and moving the chamber with respect to the Van de Graaff generator until the beam passed centrally through the entry and exit ports. The probes are shown schematically in Fig. 1 and consist of two sets of four coplanar rods with conical points. One set is located at the beam entry port and the other at the beam exit port. The probes are symmetrically located around the nominal beam position, with their pointed tips equidistant from the centerline of the port. The probes are retracted during scattering measurements so that they do not interfere with the beam. To align the beam, the currents to the four probes in the entry port are measured simultaneously and the chamber is moved laterally until the currents are equal. Then the currents to the four exit port probes are measured and the chamber is tilted until these currents are equal. The mechanism for the tilting adjustment (not shown in Fig. 1) consists of provisions for motion about a pair of axes which are perpendicular to each other and to the beam. These axes are located close to the entry port beam probes so as not to disturb the lateral alignment.

III. EXPERIMENTAL PROCEDURE

The measurements of relative scattering cross sections were made by determining the ratio of scattering at a given angle to that at 90°. The procedure is as follows. The scattering detector is placed at the angle for which the scattering is to be measured and the monitor detector is placed at the 90° position on the opposite side of the scattering chamber. The apparatus as shown in Fig. 1 is in the configuration used to measure scattering at 150°. The ratio of scattering into these two detectors is determined by counting simultaneously for a given time and recording the total number of counts. The ratio of electrons scattered into the same two detectors is then determined again with the scattering detector moved to the 90° position, directly opposite the monitor detector. By dividing the first ratio by the second, the ratio of scattering at the given angle to that at 90° is determined. This procedure is employed rather than the simpler one of merely taking the first ratio de-

[8] C. T. Chase and R. T. Cox, Phys. Rev. 58, 243 (1940).

scribed, in order to eliminate errors due to differences in collimators and detectors. This procedure also minimizes possible errors due to misalignment of the chamber with respect to the beam, since the signs of such errors would be the same for the two angles being compared. Motion of the scattering chamber with respect to the beam of as little as 0.1 mm in translation could be detected. An error of the order of 0.06 percent would result from a misalignment of this order of magnitude. It is felt that the maximum misalignment due to all causes was probably less than ten times the minimum detectable value given, resulting in a maximum probable error due to translational chamber misalignment of about 0.6 percent. The error due to possible rotational misalignment is estimated to be of the same order of magnitude as that from translational misalignment.

Each measurement consisted of three to five runs of from 20 000 to 100 000 counts for the slower scaler. To cancel out errors due to mounting the foil slightly off the axis of the foil mounting shaft, a second measurement was made for each configuration with the foil turned through 180°, and the mean of the two values was taken to be the measured scattering. In the majority of cases the two measurements did not differ by more than the statistical uncertainty of a few tenths of a percent, indicating that the foils were being centered to within a fraction of a millimeter. In all cases the foil was placed at 45° with respect to the beam.

The measurements at 30° and 60° were made with the foil in the "transmission" position, i.e., with the scattered electrons and the beam emerging from the same side of the foil, while at 120° and 150° the foil was in the "reflection" position. The measurements at 120° and 150° had to be adjusted to account for the asymmetry in the transmission-reflection scattering. This asymmetry was measured for each foil and at each energy by taking the ratio of scattering at 90° to that at 270° with the foil in the four possible positions obtainable by rotation through 90°. The square root of the quotient of the average ratios for scattering with the 90° detector measuring reflected electrons to that with the same detector measuring transmitted electrons gives the transmission-reflection asymmetry at 90°. The values obtained ranged from 3 percent for aluminum at 0.6 Mev to 0.2 percent for gold at 1.7 Mev. This measured scattering asymmetry was used to adjust the data for 120° and 150° in a manner which is discussed in the section on the treatment of data.

The pulse height discriminator setting was checked before each count was taken. As may be seen from the curves in Fig. 2, the counting rate changes very slowly with discriminator setting over a wide range of settings. The discriminator was set at a given fraction of the value at which the counting rate was half of the plateau value. Usually this fraction was 0.82 although the exact value is not important as long as the same value is used throughout a given measurement. This is true because the shape of the curve (Fig. 2) is independent of angle for pulse heights greater than about 80 percent of that representing the beam energy.

An effort was made in all scattering measurements to keep the counting rates in the two detectors within certain bounds in order to avoid large errors due to the finite resolving time of the scaling circuits and to the background from radioactivity and cosmic rays. In order to keep the corrections small, an attempt was made to have the counting rates in the two detectors differ by not much more than a factor of ten. Since the ratio of scattering at 30° to that at 90° was about 90 for aluminum, a special collimator with a solid angle about one-tenth that of the standard collimator was used at 30° in order that the counting rates would differ by no more than a factor of ten. The ratio of scattering at 150° to that at 90° was a factor of twelve and the ratios for 60° and 120° were less than this, so that standard collimators were used for all measurements except those at 30°. At 1.7 Mev, the 30° measurement was made using standard collimators by placing the monitor at the 60° position rather than at 270°.

The measurements of absolute scattering were made by comparing the number of electrons scattered at 60° with the integrated beam current. The beam charge was collected on a Plasticon condenser of 5.0-μf capacity. The potential to which the condenser was charged was measured using a vibrating reed electrometer. The scattering chamber had a long tube or cup on the bottom which acted as a Faraday cage. Approximately 90 percent of the total beam charge was collected in this cup, the remainder being collected in the scattering chamber itself, which was insulated from the Van de Graaff generator. Actual scattering measurements were made with the Faraday cup and the scattering chamber connected electrically.

Absolute scattering measurements were made using the same collimators as for the relative measurements. However the crystal was not mounted directly behind the diaphragm which defines the solid angle, as was the case in the relative measurements. The reason for this is that there was evidence to show that a significant number of electrons were being scattered from the inner edge of the defining aperture and into the detecting crystal. The crystal was mounted behind a stripping diaphragm located $1\frac{3}{4}$ in. behind the diaphragm which defines the solid angle. This stripping diaphragm had a hole which was just large enough that electrons scattered directly from the foil would not be intercepted. However those which were scattered from the edges of the beam defining diaphragm could be intercepted, since, in general, the direction of such electrons would have been changed in the process. Analysis of the results of the first absolute scattering experiment gave evidence that the precautions described were not adequate to reduce the contribution from diaphragm scattering to a negligible amount. Consequently the crystal was moved to a position 6 in. behind the beam defining diaphragm and two stripping diaphragms were

used, one next to the crystal and the other at an intermediate position.

The scatterer used in the first measurements was an aluminum foil positioned perpendicular to the beam and intersecting it at the center of the scattering chamber. The foil was mounted in such a way that it could be shifted parallel to itself between measurements without disturbing the chamber alignment. Four positions were chosen for the foil mount, corresponding to four spots in the foil for which scattering measurements could be made. These same four positions were used for the measurements at the three energies. After the scattering measurements were made, the foil was cut into two rectangles approximately 1.8 cm×2.4 cm, each containing two of the scattering spots. The dimensions and the masses of the rectangular foils were measured to determine their average thicknesses.

Another determination of absolute scattering was made at 1.7 Mev using the two stripping diaphragms previously described. Four separate foils were used in this measurement. After the absolute scattering was determined, pieces approximately 1.6-cm-square were cut from the foils with the scattering spot in the center. The thicknesses of these squares were then determined in the same way as in the first experiment.

IV. TREATMENT OF DATA

In analyzing the data, account was taken of background counting rate, counting losses due to the finite resolving time of the scaling circuits, multiple and plural scattering, and scattering in the Parlodion foil support. The maximum contribution for each of these was of the order of a percent and was in most cases less than this.

The background counting rate due to cosmic rays and other natural sources was determined for each set of data by taking counts for 300 to 500 sec with the Van de Graaff generator off. The variations in the background counting rate at a given energy were no more than expected from statistical fluctuations. The average background counting rate was about 1.5 counts/sec for 0.6 Mev, 0.3 counts/sec for 1 Mev, and 0.1 counts/sec for 1.7 Mev. The background counting rate was never more than about one percent of the total counting rate.

The background due to x-rays and to stray electrons not originally scattered by the foil was determined by placing a blank foil holder in the scattering chamber. This was done by mounting two foil frames, one with a foil and one without, on a rotating plate arranged so that either frame could be positioned at the center of the chamber. The apparatus was first set up using the foil, and then the blank frame was placed in the scattering position. Measurements were made at 1 Mev using an aluminum foil. The counting rate with the blank foil frame in place represents the sum of the natural background and the background due to x-rays and stray electrons, while the counting rate with the Van de Graaff off represents only the natural background. The maximum contribution to the background from the Van de Graaff generator itself amounted to about 0.01 percent of a typical scattering measurement, and was therefore neglected in analyzing the data.

The measured counting rates were adjusted for losses due to the finite resolving time of the scaling circuit. For low enough counting rates this loss is directly proportional to the counting rate. The resolving time of the circuit was measured by determining the difference in counts recorded at different counting rates for a given integrated beam current. The value determined in this way, 5.6 ± 0.5 μsec, compares with the value of ~ 6 μsec determined with a double pulse generator.

In taking account of multiple scattering, an equation derived by Chase and Cox[8] was used. This equation is based on the assumption that the electrons are deflected through a large angle θ' in a single scattering event and through a small angle ϵ by multiple scattering in such a way that the resultant is the experimental

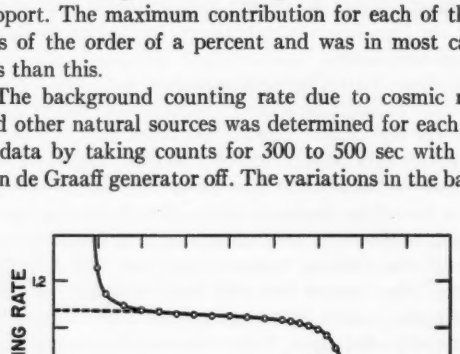


FIG. 3. Integral pulse height distribution obtained in absolute scattering measurement. Dashed portion of curve shows extrapolation of plateau.

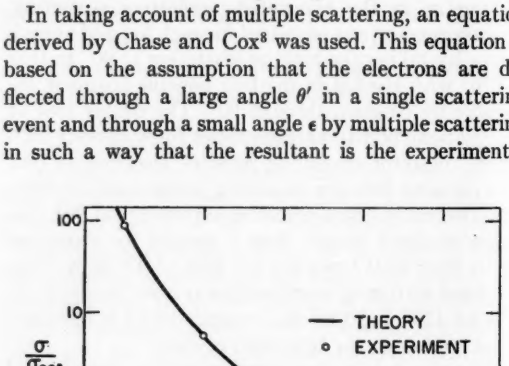


FIG. 4. Differential scattering cross section of aluminum for 1-Mev electrons. Measurements are normalized at 90°.

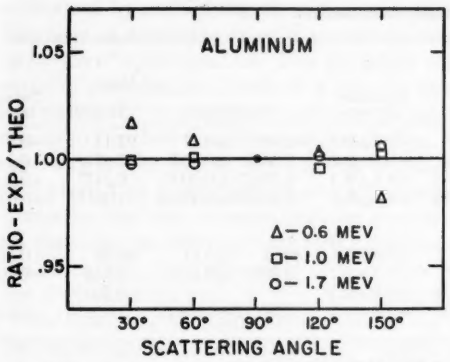


FIG. 5. Differential scattering cross sections of aluminum for 0.6-, 1.0-, and 1.7-Mev electrons, expressed as the ratio of experimental to theoretical values. Measurements are normalized at 90°.

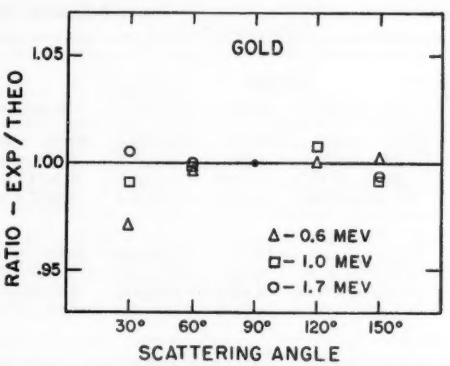


FIG. 6. Differential scattering cross sections of gold for 0.6-, 1.0- and 1.7-Mev electrons, expressed as the ratio of experimental to theoretical values. Measurements are normalized at 90°.

angle θ of interest. Their expression is

$$\frac{P'(\theta)-P(\theta)}{P(\theta)}=\epsilon^2\left(\csc^2\frac{\theta}{2}-\frac{1}{2}\right),$$

where $P(\theta)$ is the probability of a single scattering through an angle θ and $P'(\theta)$ is the probability of a single scattering through θ' and multiple scattering through ϵ resulting in a net scattering through θ, averaged over all possible values of θ'. For the foil thicknesses used this correction was negligible at all angles except 30°, where it was of the order of one to three percent. The mean square multiple scattering angle $\langle\epsilon^2\rangle_{Av}$ was obtained by assuming that the angular distribution of multiple scattered electrons was gaussian with a distribution $\exp(-\epsilon^2/2\langle\epsilon^2\rangle_{Av})$ with $\langle\epsilon^2\rangle_{Av}$ given by

$$\langle\epsilon^2\rangle_{Av}=\frac{4\pi Nxe^4Z(Z+1)}{p^2v^2}G.$$

Here e, p, and v are the charge, momentum, and velocity of the electron, N the number of nuclei per unit volume, Z the atomic number of the scatterer, x the foil thickness, and G a very slowly varying function of Z and p.[9] The value of G was measured experimentally by Kinzinger and Bothe[3] for the case of 245-kev electrons scattered in aluminum. Their value of $G=2.3$ was used in adjusting the data of all measurements, since G is for practical purposes a constant. The contributions from multiple scattering were small enough that the choice of G was relatively unimportant. The value of $(\langle\epsilon^2\rangle_{Av})^{\frac{1}{2}}$ was of the order of three degrees or smaller for all foils and energies used.

The effects of the divergence of the focused beam and of the spread in angle of the measured electrons due to the finite size of the collimator aperture are more important at 30° than at any of the other angles for which measurements were made. The divergence of the beam is largest for 0.6 Mev and is estimated to be no larger than 0.6° (from nominal beam position to the electron path of maximum deviation). If the divergence were actually this large, the scattering at 30° would be only 0.15 percent greater than for a truly parallel beam. For other angles and energies the effect would be even less, so this effect was neglected.

The effect on the measured values of scattering due to the fact that the measurements extended over a small angular range centered about the angle of interest, is most pronounced at small scattering angles. However at 30° for 0.6 and 1.0 Mev, a collimator was used which had a considerably smaller angular spread than the standard collimator. At 1.7 Mev, where the standard collimator was used at 30°, the adjustment required was only 0.1 percent. This adjustment was the only one made since for all other measurements they were considerably smaller than 0.1 percent.

Plural scattering, wherein a small number of scattering events results in an electron being deflected through the angle of interest, is quite small for the foil thicknesses used in the experiment. The most probable event of this type is one in which an electron is first scattered into the plane of the foil and is then scattered out of the foil in the direction of interest. The path length of the electron through the foil is much larger in this case than for the single scattering, and the probability for this double scattering is correspondingly higher than would be expected for a simple double scattering in the foil. For the foils used in this experiment the effect was measured by placing detectors at the 90° and the 270° positions and comparing the scattering from the "reflection" side of the foil with that from the "transmission" side. To eliminate differences in detectors, measurements were made with the foil oriented so that first one detector faces the transmission side of the foil and then the other. The asymmetry thus obtained was considered to be due entirely to scattering twice through 45° to give 90°. Adjustments were made in the scattering measurements in which the electrons were scattered from the incident or reflection side of the foil, namely

[9] Groetzinger, Berger, and Ribe, Phys. Rev. 77, 584 (1950).

TABLE I. Relative differential scattering cross section.

Angle	Aluminum 0.6 Mev Experiment[a]	Theory	Deviation percent[a]	1.0 Mev Experiment[a]	Theory	Deviation percent[a]	1.7 Mev Experiment[a]	Theory	Deviation percent[a]
30°	84.67 ±0.23	83.31	+1.6±0.3	89.06 ±0.27	89.08	−0.0±0.3	92.99 ±0.27	93.15	−0.2±0.3
60°	5.291 ±0.011	5.244	+0.9±0.2	5.504 ±0.010	5.496	+0.1±0.2	5.662 ±0.017	5.674	−0.2±0.3
120°	0.2969±0.0007	0.2960	+0.3±0.3	0.2654±0.0007	0.2665	−0.4±0.3	0.2462 ±0.0005	0.2457	+0.2±0.2
150°	0.1166±0.0003	0.1188	−1.8±0.3	0.0857±0.0003	0.0854	+0.3±0.3	0.06221±0.00018	0.06187	+0.6±0.3
	Gold								
30°	37.66 ±0.10	38.79	−2.9±0.3	39.50 ±0.12	39.87	−0.9±0.3	40.71 ±0.12	40.49	+0.5±0.3
60°	3.767 ±0.010	3.781	−0.4±0.2	3.898 ±0.009	3.889	−0.2±0.2	3.950 ±0.011	3.954	−0.1±0.3
120°	0.3403±0.0010	0.3404	−0.0±0.2	0.3204±0.0009	0.3178	+0.8±0.3		0.3036	
150°	0.1296±0.0004	0.1294	+0.2±0.3	0.0961±0.0004	0.0969	−0.8±0.4	0.07612±0.00022	0.07656	−0.6±0.3

[a] The standard deviation due to counting statistics is given with the experimental values and with the deviations.

those at 120° and 150°. The asymmetries at these two angles were determined by assuming that electrons were scattered into the plane of the foil (45°) and then through 75° and 105° to give a total of 120° and 150° respectively. The experimentally measured 45°–45° asymmetry was multiplied by the theoretical ratios of scattering at 75°/45° and 105°/45° to give asymmetries for 45°–75° and 45°–105° double scattering.

The scattering measurements in gold were adjusted for the scattering in the Parlodion foil supports by measuring directly the contribution of the Parlodion at 60°. These measurements were made by determining the absolute scattering from the foil support before and after the gold was evaporated on. The contributions at angles other than 60° were calculated by normalizing at the measured 60° value and using the theoretical cross sections for Parlodion. To check the validity of using the theoretical cross sections for Parlodion, a measurement of the scattering from a Parlodion foil was made which agreed with the theory within the experimental error. Since the Parlodion consists mostly of hydrogen and carbon, and since the relative scattering is an insensitive function of Z for small Z, the form of the relative scattering in the foil support was essentially the same as in the aluminum. Consequently the effect of the Parlodion backing was insignificant in the case of aluminum and no adjustments were made.

The integral bias curve for the absolute scattering measurements, shown in Fig. 3, does not have a perfectly flat "plateau" but the counting rate increases slowly as the pulse height decreases. These data were taken using the improved detector geometry. The rapid rise at small pulse heights is caused mostly by x-rays from the beam electrons, which are stopped in the Faraday cage and in the chamber. The slope of the plateau was considered to be due principally to electrons having been backscattered out of the anthracene crystal before dissipating all of their energy. The energy distribution of electrons backscattered from carbon[10] was used in extrapolating the plateau back to zero pulse height. In this extrapolation, the energy distribution was normalized by assuming that the slope of the plateau between 40 and 60 percent of the beam energy was due entirely to backscattering from the crystal.

[10] W. Bothe, Z. Naturforsch. 4a, 542 (1949).

V. RESULTS

The derivation given by Mott[1] for the coulomb scattering of electrons is based on Dirac's relativistic quantum mechanics. The resulting differential scattering cross section is

$$d\sigma/d\Omega = q^2(1-\beta^2)FF^* \csc^2(\theta/2) + GG^* \sec^2(\theta/2),$$

where

$$q = \alpha/\beta = Ze^2/\beta\hbar c.$$

Here β is the electron velocity in units of the velocity of light, and F and G are infinite series in the Legendre polynomials of $\cos\theta$. The cross section is in units of $(\hbar/p)^2$ where p is the momentum of the electron.

The differential cross section for the scattering of electrons by Hg ($Z=80$) has been evaluated numerically by Bartlett and Watson.[11] They give the results for various energies up to ~1.7 Mev and for several scattering angles from 15° to 180°. McKinley and Feshbach[12] have expanded Mott's series for F and G as a power series in $\alpha(=Ze^2/\hbar c \cong Z/137)$ and α/β and have calculated the coefficients of terms to fourth order in α. Assuming that the coefficients of the terms in α^5 and higher are not $\gg 1$, the errors in the scattering calculations for aluminum caused by neglect of terms in α^5 should be considerably smaller than the experimental errors, since α^5 for aluminum is 8×10^{-6}. The assumption that the coefficients of the higher order terms are not $\gg 1$ is plausible because of the fact that the calculated coefficients are all of the order of unity or smaller. In contrast with the case for aluminum, this α^4 approximation for gold is not expected to yield accurate results since $\alpha^5 = 0.06$. The experimental measurements in gold were compared with values obtained from those given by Bartlett and Watson. The cross sections for gold

[11] J. H. Bartlett and R. E. Watson, Proc. Am. Acad. Arts Sci. 74, 53 (1940).

[12] W. A. McKinley, Jr. and H. Feshbach, Phys. Rev. 74, 1759 (1948).

were obtained by multiplying these values for mercury by the ratio of scattering in gold to that in mercury given by the α^4 approximation. This procedure may be expected to give results accurate to within one percent.

The results for the relative scattering in gold and aluminum at the three energies are shown in Figs. 4 to 6 and in Table I. The data in Fig. 4 are for aluminum at 1 Mev, plotted on a semi-logarithmic scale. Since the diameters of the dots representing the experimental points are about ten times as large as the actual experimental deviations, it is impossible to show the extent of the deviations on such a plot. Consequently, in Fig. 5 the same data are plotted in the form of the ratio of the experimental to the theoretical values. The data for aluminum at 0.6 Mev and at 1.7 Mev are also shown in Fig. 5. The experimental values for 1.7 Mev agree with the theory as shown. The results at 0.6 Mev do not agree quite as well as those at 1.7 Mev, but the maximum deviation is only 1.8 percent.

The results for gold are shown in Fig. 6. Here the results deviate from the theory a little more than was the case for aluminum. However, there is no evidence for a consistent trend in the deviation. The measured value at 30° and 0.6 Mev is 2.9 percent less than the theoretical value. The cause of this deviation is not known but because of the consistent agreement of other measurements it is quite likely to be an instrumental error. Before adjustments were made in this datum for scattering in the Parlodion foil support and for multiple scattering, the measured value was 0.7 percent greater than the theoretical value. The measurement of scattering at 120° and 1.7 Mev was not completed because of Van de Graaff generator failure.

The results of the measurements on absolute scattering are given in Table II. The results at 0.6 Mev and 1.0 Mev agree with the theory within the standard deviations of the four determinations made at each energy. The first measurement at 1.7 Mev gives results which are somewhat higher than the theoretical value. It can be seen that the initial measurements (the first three tabulated in Table II) show an increase in experimental cross section with increasing energy, when compared with the theory. An analysis of the results indicated that the stripping diaphragm was probably not intercepting a majority of the electrons at the high energies. Consequently another experiment was performed at 1.7 Mev with the improved stripping diaphragm geometry described previously. The slope of the plateau of the integral bias curve for this second experiment was reduced to 54 percent of the previous value. This indicates that the first stripping diaphragm geometry had indeed not been completely effective. The results of this second measurement at 1.7 Mev are given in Table II. The larger standard deviation given for the second experiment probably results from the fact that the foil pieces were smaller and hence their thicknesses were more difficult to determine than was the case in the first experiment. Despite the larger standard deviation, the agreement with theory is better for the second determination than for the first, and is well within this larger standard deviation.

TABLE II. Absolute differential scattering cross section for aluminum at 60°.

Energy Mev	Experiment 10^{-24} cm²[a]	Theory 10^{-24} cm²	Deviation percent[a]
0.6[b]	15.95±0.27[b]	16.08	−0.8±1.7[b]
1.0[b]	6.84±0.10[b]	6.76	+1.2±1.5[b]
1.7[b]	2.84±0.041[b]	2.705	+4.8±1.4[b]
1.7[c]	2.80±0.15[c]	2.705	+3.4±5.4[c]

[a] The standard deviation of the four measurements made is given with the experimental value and with the deviation.
[b] Measurements made with original slit-edge-scattering stripping geometry, 1½ in. long with a single stripping diaphragm.
[c] Measurement made with improved slit-edge-scattering stripping geometry, 6 in. long with two stripping diaphragms.

VI. ACKNOWLEDGMENTS

The authors wish to express their appreciation for the contributions made by A. G. McElroy, T. R. Murray, B. A. Hess, and Mrs. Helen Connelley.

Theory of Hyperfine Structure*

CHARLES SCHWARTZ†
Department of Physics and Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Massachusetts
(Received September 24, 1954)

Considering the classical electric and magnetic interactions between atomic electrons and the nucleus, we arrive at a representation of the hyperfine interactions in terms of a multipole expansion of the field potentials. Treating these noncentral interactions in first order perturbation theory we can give the form of the general interval structure and analyze for the multipole interaction constants using Racah coefficients. Pertinent matrix elements for a single valence electron are calculated relativistically. Some second order terms of the dipole and quadrupole interactions are calculated as they affect the interpretation of the first order octupole interaction. In this work we also take into account quantitatively the effect of some electronic configuration interaction. Finally the values of nuclear magnetic octupole moments expected according to different models are calculated and compared with the experimental data thus far collected. Generally the measured octupole moments are in as good agreement with the values predicted by the single-particle shell model as are the corresponding dipole moments.

INTRODUCTION

THE recent very accurate measurements by Jaccarino *et al.*[1] on the hyperfine structure of the ground state of I^{127} showed that the theory of dipole and quadrupole interactions as previously developed was insufficient to describe the level structure to this new high precision. In order to learn how finer details of the level structure could be interpreted in terms of higher nuclear moments, it was decided to review the entire theory of hyperfine structure.

First, treating the nucleus as a stationary nonrelativistic quantum mechanical system, we derive a multipole expansion for the electric and magnetic fields produced by the nuclear charge and current distributions. In this development the electromagnetic potentials are expressed in terms of scalar and vector spherical harmonics. When these potentials are put into the Dirac Hamiltonian for the electrons, the terms of different orders of symmetry can be identified as tensor operators, so that one can write down the form of the generalized "interval rule" in terms of Racah coefficients. This analysis proves to have not only formal, but also practical computational advantages over the formulations given in the prior literature.

The interaction constant A_k for each multipole order k is the product of the nuclear moment of that order and an electronic matrix element. The general electronic matrix elements are evaluated for the case of a single valence electron using the techniques of Racah[2] for the spin-angular integrals, and following the method of Casimir,[3] Racah,[4] and Breit[5] for the radial integrals.

The magnitude of the hyperfine interaction energies decreases rapidly with order: the magnetic octupole interaction is weaker than the magnetic dipole by about 10^{-5}. Thus in the perturbation theory second order terms (mixing in excited electron states) in the dipole and quadrupole interactions give contributions which may appear as first order magnetic octupole (and electric 2^4-pole) interactions. This effect is calculated using so far only the perturbation of the nearby doublet level. In this work the dipole contribution of s-electrons in mixed configurations is allowed for by quantitative analysis depending on the measured dipole interaction constants in both states of the doublet.

After all the electronic contributions have been extracted from the observed octupole interaction constant the value of the nuclear magnetic octupole moment is finally revealed. We have calculated the values of these moments to be expected according to the individual-particle shell model for the nucleus. The results, for various single-particle orbits, are represented on a diagram similar to the Schmidt plot for dipole moments; and the octupole moments of the few nuclei already investigated take approximately the same positions on this new diagram as they do on the Schmidt plot.

The measurement of higher nuclear moments may also prove to be a valuable test of the Bohr-Mottelson collective nuclear model. In their "strong-coupling" scheme the observable magnitude of an octupole moment is decreased from the single-particle value of the simpler theory by a factor as small as 1/35.

THE HYPERFINE STRUCTURE

In the absence of external fields an atom may be described very accurately in terms of a compact charged nucleus and an electron system arranged in the central coulomb field produced by the nucleus. If, however, the electrons have some resultant angular momentum $J>0$ and the nucleus also has a spin $I>0$, there will be further interactions between the two

* This work was supported in part by the Signal Corps; The Office of Scientific Research, the Air Research & Development Command; and the Office of Naval Research.
† Gulf Oil Corporation Fellow 1953–54.
[1] Jaccarino, King, Satten, and Stroke, Phys. Rev. 94, 1798 (1954).
[2] G. Racah, Phys. Rev. 62, 438 (1942).
[3] H. B. G. Casimir, *On the Interaction Between Atomic Nuclei and Electrons* (Teyler's Tweede Genootschap, Haarlem, 1936).
[4] G. Racah, Nuovo cimento 8, 178 (1931).
[5] G. Breit, Phys. Rev. 38, 463 (1931); and G. Breit and W. A. Wills, Phys. Rev. 44, 470 (1933).

systems, described as magnetic dipole, electric quadrupole, etc. We then denote by hyperfine structure (hfs) the different energy eigenvalues associated with the various total angular momentum states of the combined systems characterized by the quantum number F:

$$|I-J| \leq F \leq I+J.$$

The major part of the electron-nuclear (noncentral) interaction is the magnetic dipole term which is characteristically of the order of magnitude

$$\mu_0\mu_N\left\langle\frac{1}{r_e^3}\right\rangle \sim \frac{e\hbar}{2mc}\frac{e\hbar}{2Mc}\left\langle\frac{1}{r_e^3}\right\rangle,$$

where m is the electron mass, M the proton mass, and r_e the radial coordinate of the electron from the nucleus. This may be compared with the fine-structure (fs) spacing of the electronic energy levels, which is of the order

$$(\hbar/2mc)^2Ze^2\langle 1/r_e^3\rangle,$$

where Z is the atomic number. The ratio of these two

$$\text{hfs/fs} \sim m/MZ \sim 10^{-4}-10^{-5},$$

is then a measure of the approximation to which we may use the various orders of perturbation theory to calculate the hfs levels. We shall start by considering the hfs interactions in first order and later turn our attention to the effect of second order terms on the higher moments.

First, the noncentral interactions between electrons and the nuclear particles, whatever these interactions may be, will be expanded in a series of tensor operators. The perturbation Hamiltonian H_1 is written

$$H_1=\sum_k T_e^{(k)}\cdot T_n^{(k)}. \tag{1}$$

$T_e^{(k)}$ is a tensor operator of rank k which operates in the space of the electronic coordinates only; its rank is defined by the fact that it commutes with the total angular momentum operator of the same space, $\mathbf{J}$, just as do the spherical harmonics of order k. $T_n^{(k)}$ operates on the coordinates of the nucleons in the same manner; and the terms in the series (1) are the scalar products of these two tensors, thus are invariants in the combined space.[6] We now wish to calculate the first order energy expectation values of H_1 in states described as having J of the electron and I of the nucleus coupled to the resultant F.

$$W_F=\langle IJF|H_1|IJF\rangle=\sum_k\langle IJF|T_e^{(k)}\cdot T_n^{(k)}|IJF\rangle. \tag{2}$$

These matrix elements are independent of the magnetic quantum number M_F, so it will be ignored. According to a well-known theorem of Racah,[2] the dependence on F of each of the matrix elements in (2) can be separated out as follows:

$$\langle IJF|T_e^{(k)}\cdot T_n^{(k)}|IJF\rangle=(-1)^{I+J-F}W(IJIJ;Fk)\times\langle I\|T_n^{(k)}\|I\rangle\langle J\|T_e^{(k)}\|J\rangle,$$

where W, the Racah coefficient, is a known algebraic function of its six arguments; and the double-barred matrix elements, called the reduced matrix elements, are independent of any magnetic quantum numbers which may be assigned to the states indicated.

We shall write the hfs term energy as

$$W_F=\sum_k A_kM(IJ;F;k), \tag{3a}$$

with the normalization

$$M(IJ;I+J;k)=1, \tag{3b}$$

which gives directly

$$A_k=\langle T_e^{(k)}\rangle_{JJ}\langle T_n^{(k)}\rangle_{II}, \tag{4a}$$

with the relation

$$\langle T^{(k)}\rangle_{JJ}=\frac{(2J)!}{[(2J-k)!(2J+k+1)!]^{\frac{1}{2}}}\langle J\|T^{(k)}\|J\rangle. \tag{4b}$$

Our A's are related to the usual[1] hfs interaction constants as follows:

$$A_1=IJa, \quad A_2=1/4b, \quad A_3=c. \tag{4c}$$

The coefficient M is given by the formula:

$$M(IJ;F;k)=\frac{(2I-k)!}{(2I)!}\frac{(2J-k)!}{(2J)!}\frac{(I+J-F)!(J-I+F)!(I-J+F)!}{(I+J+F+1)!}[k!]^2\times\sum_z(-1)^{z+I+J-F}\frac{(2I+2J+1-z)!}{z!(2I-k-z)!(2J-k-z)![(I+J-F-z)!(k+F-I-J+z)!]^2}, \tag{5}$$

where the sum extends over all integral values of z for which no factorial has a negative argument.

It has been customary to express M in terms of the cosine factor,

$$K=F(F+1)-I(I+1)-J(J+1).$$

[6] The heavy dot $\cdot$ will be used to denote the scalar product of two tensor operators,

$$T^{(k)}\cdot U^{(k)}=\Sigma_\mu T_\mu^{(k)}U_{-\mu}^{(k)}(-1)^\mu,$$

while the light dot $\cdot$ will denote the scalar product of two cartesian vectors,

$$\mathbf{V}\cdot\mathbf{W}=V_xW_x+V_yW_y+V_zW_z.$$

The formulas for the first four orders of M in terms of K are the following‡:

$$M(IJ;F;1)=\frac{1}{2IJ}K, \tag{6a}$$

$$M(IJ;F;2)=\frac{6}{(2I)(2I-1)(2J)(2J-1)}[K(K+1)-(4/3)I(I+1)J(J+1)], \tag{6b}$$

$$M(IJ;F;3)=\frac{20}{(2I)(2I-1)(2I-2)(2J)(2J-1)(2J-2)}[K^3+4K^2$$
$$+(4/5)K\{-3I(I+1)J(J+1)+I(I+1)+J(J+1)+3\}-4I(I+1)J(J+1)], \tag{6c}$$

$$M(IJ;F;4)=\frac{70}{(2I)(2I-1)(2I-2)(2I-3)(2J)(2J-1)(2J-2)(2J-3)}[K^4+10K^3$$
$$+(4/7)K^2\{-6I(I+1)J(J+1)+5I(I+1)+5J(J+1)+39\}$$
$$+(4/7)K\{-34I(I+1)J(J+1)+12I(I+1)+12J(J+1)+18\}$$
$$-(24/35)I(I+1)J(J+1)\{-2I(I+1)J(J+1)+4I(I+1)+4J(J+1)+27\}]. \tag{6d}$$

It should be pointed out that formulas (6c, d) are quite unwieldy for numerical evaluation and it is frequently easier to work directly from (5). For example, if $2J=k$ we have

$$M(IJ;F;2J)=(-1)^{I+J-F}$$
$$\times\frac{(2J)!(2I+2J+1)!(I-J+F)!}{(2I)!(I+J+F+1)!(I+J-F)!(J-I+F)!}, \tag{7a}$$

or if $k=2J-1$,

$$M(IJ;F;2J-1)=(-1)^{I+J-F}$$
$$\times\frac{(2J-1)!(2I+2J)!(I-J+F)!}{(2I)!(I+J+F+1)!(I+J-F)!(J-I+F)!}$$
$$\times 2[F(F+1)-I(I+1)+J^2]. \tag{7b}$$

The following sum rule is also of help in checking numerical work

$$\sum_F(2F+1)M(IJ;F;k)=0 \quad (k>0).$$

Aside from the physical content of the operators (1), this analysis gives us the selection rule that the series (3a) terminates at the term $k=2J$ or $2I$, whichever is smaller. One then has $2J$ (or $2I$) interaction constants A_k to be solved for from the $2J$ (or $2I$) measured energy intervals. Because of an orthogonality sum of the Racah coefficients one can solve (3a) analytically for the A_k.

$$A_k=(2k+1)\frac{[(2I)!(2J)!]^2}{(2I-k)!(2I+k+1)!(2J-k)!(2J+k+1)!}$$
$$\times\sum_F(2F+1)M(IJ;F;k)W_F. \tag{8}$$

ELECTROMAGNETIC POTENTIALS

We shall now describe the electric and magnetic static interactions between the nuclear and electronic systems in a multipole expansion.

The electrostatic potential set up in space by a distribution of charges in the nucleus is

$$V(r)=\int\frac{\rho(r')}{|r-r'|}dv', \tag{9}$$

where $\rho=e\Psi^*\sum_i g_{li}\Psi$ is the density of electric charge of all the nucleons; g_l is $+1$ for protons and zero for neutrons. Now, expanding the Green's function in terms of spherical harmonics,

$$\frac{1}{|r-r'|}=\sum_k r_<^k r_>^{-k-1}C^{(k)}(\theta,\varphi)\cdot C^{(k)}(\theta',\varphi'), \tag{10}$$

where

$$C_\mu^{(k)}(\theta,\varphi)=\left(\frac{4\pi}{2k+1}\right)^{\frac{1}{2}}Y_{k\mu}(\theta,\varphi),$$

we have the desired multipole operators of the electric interaction. The functions $C^{(k)}$ are tensor operators of rank k with parity $(-1)^k$.

The vector potential for the magnetic field set up by the nuclear currents and spins is not quite so simple; it will be expressed in terms of vector spherical harmonics.[7] We choose the form

$$\mathbf{A}=\sum_{k\mu}\mathbf{L}C_\mu^{(k)}(\theta,\varphi)h_{k,\mu}(r), \tag{11}$$

‡ Formula (6c) has been derived previously by Kramers [Proc. Roy. Acad. Amsterdam 34, 965 (1931)] and Nierenberg [Phys. Rev. 87, 225 (1952)]. See also Biedenharn, Blatt, and Rose [Revs. Modern Phys. 24, 249 (1952)] for a general recursion formula for these coefficients.

[7] J. M. Blatt and V. F. Weisskopf, *Theoretical Nuclear Physics* (John Wiley and Sons, Inc., New York, 1952), Appendix B, p. 796.

where $\mathbf{L}=-i\mathbf{r}\times\text{grad}$ which assures the gauge $\text{div}\mathbf{A}=0$. The operator $\mathbf{L}$ commutes with any function of r and also with the Laplacian so that the equation $\nabla^2\mathbf{A}=-(4\pi/c)\mathbf{j}$ becomes

$$\sum_{k\mu}\mathbf{L}C_\mu^{(k)}(\theta,\varphi)\left[\frac{1}{r}\frac{d^2}{dr^2}r-\frac{k(k+1)}{r^2}\right]h_{k,\mu}(r)=-\frac{4\pi}{c}\mathbf{j}. \quad (12)$$

Now using the orthogonality of the vector spherical harmonics over the unit sphere,

$$\int[\mathbf{L}C_\mu^{(k)}]^*\cdot[\mathbf{L}C_{\mu'}^{(k')}]d\Omega=\delta_{kk'}\delta_{\mu\mu'}\frac{4\pi}{2k+1}k(k+1), \quad (13)$$

we get

$$\left[\frac{1}{r}\frac{d^2}{dr^2}r-\frac{k(k+1)}{r^2}\right]h_{k,\mu}(r)\cdot$$

$$=-\frac{1}{c}\frac{2k+1}{k(k+1)}\int[\mathbf{L}C_\mu^{(k)}(\theta,\varphi)]^*\cdot\mathbf{j}d\Omega.$$

The Green's function for the left-hand side is

$$\frac{1}{2k+1}r_<{}^k r_>{}^{-k-1},$$

so finally we have the solution

$$h_{k,\mu}(r)=-\frac{1}{c}\frac{1}{k(k+1)}\int[\mathbf{L}r_<{}^k r_>{}^{-k-1}C_\mu^{(k)}(\theta,\varphi)]^*\cdot\mathbf{j}dv. \quad (14)$$

The nuclear currents $\mathbf{j}$ consist of two parts: the convection current

$$\mathbf{j}_c=e\Psi^*\textstyle\sum_i g_{li}\mathbf{v}_i\Psi,$$

and the spin current

$$\mathbf{j}_s=\text{curl}\Psi^*\textstyle\sum_i g_{si}\dfrac{e\hbar}{2m_i}\mathbf{S}_i\Psi.$$

The convection current should be a symmetrized combination but it is easily shown that the two terms give just the same contributions under the integral (14).

For the convection current term in (14) we can write

$$[\mathbf{L}(\ \ )]^*\cdot\mathbf{j}_c=i\mathbf{r}\times\nabla(\ \ )^*\cdot\mathbf{j}_c=-i\nabla(\ \ )^*\cdot(\mathbf{r}\times\mathbf{j}_c),$$

and get its contribution to (14) in the convenient form

$$\frac{-i}{k(k+1)}\int[\nabla r_<{}^k r_>{}^{-k-1}C_\mu^{(k)*}(\theta,\varphi)]\cdot\Psi^*2g_l\mu_N\mathbf{L}\Psi dv, \quad (15)$$

with the shorthand notation:

$$2g_l\mu_N\mathbf{L}\equiv\textstyle\sum_i g_{li}\dfrac{e\hbar}{m_ic}\mathbf{L}_i.$$

For the spin current term we use the identity

$$\mathbf{r}\times\text{curl}=\text{grad}\mathbf{r}\cdot-(r\partial/\partial r+1)$$

and by partial integrations we have the spin contribution to (14).[8]

$$\frac{-i}{k(k+1)}\int\left[\nabla\left(r\frac{\partial}{\partial r}+1\right)r^k{}_<r_>{}^{-k-1}C_\mu^{(k)*}(\theta,\varphi)\right]$$

$$\cdot\Psi^*g_s\mu_N\mathbf{S}\Psi dv, \quad (16)$$

where again

$$g_s\mu_N\mathbf{S}\equiv\textstyle\sum_i g_{si}\dfrac{e\hbar}{2m_ic}\mathbf{S}_i.$$

Finally we can write the solution for the vector potential

$$\mathbf{A}=\sum_k\frac{-i\mu_N}{k(k+1)}[\mathbf{L}C^{(k)}(\theta,\varphi)]$$

$$\cdot\left\{r^{-k-1}\int_0^r\Psi^*[\nabla r'^kC^{(k)}(\theta',\varphi')\right.$$

$$\cdot(2g_l\mathbf{L}+(k+1)g_s\mathbf{S})]\Psi dv'$$

$$+r^k\int_r^\infty\Psi^*[\nabla r'^{-k-1}C^{(k)}(\theta',\varphi')$$

$$\left.\cdot(2g_l\mathbf{L}-kg_s\mathbf{S})]\Psi dv'\right\}. \quad (17)$$

In what follows we shall consider the nucleus as a point source and use only that part of the potentials corresponding to an observation point outside the nuclear matter. The error made this way affects only the magnitude of interaction observed, not its multipolarity. This error, involved in the hfs anomaly[9] is appreciable only for 2^k-pole magnetic interaction with an electron in a state $j=k/2$, and then for various orders the effect varies as $1/(k+1)$; for the dipole $(k=1)$ this effect is only a few percent in the heaviest atoms.

We can now define the integrals occurring in (9) and (17) as the general nuclear electric and magnetic multipole moments:

$$Q_k{}^\mu=e\int\Psi^*g_lr^kC_\mu^{(k)}(\theta,\varphi)\Psi dv, \quad (18a)$$

$$M_k{}^\mu=\mu_N\int\Psi^*(\nabla r^kC_\mu^{(k)}(\theta,\varphi))$$

$$\cdot\left(g_l\frac{2}{k+1}\mathbf{L}+g_s\mathbf{S}\right)\Psi dv, \quad (18b)$$

[8] There are surface terms left over from these partial integrations but they exactly cancel each other.

[9] A. Bohr and V. F. Weisskopf, Phys. Rev. 77, 94 (1950).

and write

$$V=\sum_k r^{-k-1}C^{(k)}(\theta,\varphi)\cdot Q_k, \tag{19a}$$

$$\mathbf{A}=\sum_k \frac{-i}{k} r^{-k-1}(\mathbf{L}C^{(k)}(\theta,\varphi))\cdot M_k. \tag{19b}$$

We also note that the operator Q_k has parity $(-1)^k$ while M_k has parity $(-1)^{k+1}$ so that the only static electric moments are of even order and the only static magnetic moments are of odd order if we assume that the nuclear wave function has a well-defined parity.

ELECTRONIC MATRIX ELEMENTS

We now investigate the interaction of a single electron (charge $-e$) with the nuclear fields just described. The electron wave function Ψ obeys the Dirac equation,

$$[\boldsymbol{\alpha}\cdot(c\mathbf{p}+e\mathbf{A})+\beta mc^2-e\mathrm{V}]\Psi=E\Psi, \tag{20}$$

and for the zero order solutions in only the central part of the electrostatic potential, V_c, we put $\Psi=\begin{pmatrix}\psi\\ \phi\end{pmatrix}$, where ψ is the large component and ϕ the small component of the four-spinor Ψ. Now, introducing quantum numbers, we have the separations

$$\psi_{ljm}=f(r)/r\,\mathcal{Y}_{ljm},\quad \phi_{ljm}=g(r)/ir\,\mathcal{Y}_{\bar{l}jm},$$

where

$$\bar{l}=l\pm1,\quad j=l\pm\tfrac{1}{2}=\bar{l}\mp\tfrac{1}{2}$$

$$\mathcal{Y}_{ljm}=\sum_{m_l m_s}(\tfrac{1}{2}m_s l m_l|\tfrac{1}{2}ljm)Y_{lm_l}(\theta,\varphi)\chi_{\frac{1}{2}m_s}, \tag{21}$$

and χ is a two-component spinor.

The interaction Hamiltonian is

$$H_1=-e(V-V_c)+e\boldsymbol{\alpha}\cdot\mathbf{A}, \tag{22}$$

and we will need the general matrix elements

$$\int\Psi^*H_1\Psi'dv.$$

For the electric 2^k-pole matrix element ($k>0$), these are

$$-e\langle Q_k\rangle\cdot\int_0^\infty r^{-k-1}dr[ff'(\tfrac{1}{2}ljm|C^{(k)}|\tfrac{1}{2}l'j'm')+gg'(\tfrac{1}{2}\bar{l}jm|C^{(k)}|\tfrac{1}{2}\bar{l}'j'm')].$$

Now the matrix elements of $C^{(k)}$ in the $\frac{1}{2}ljm$ scheme are independent of the quantum numbers l except for the parity selection ($l+l'+k$ even). Hence for the reduced matrix elements of electric multipoles we have

$$-e\langle\|Q_k\|\rangle\int_0^\infty r^{-k-1}(ff'+gg')dr(\tfrac{1}{2}lj\|C^{(k)}\|\tfrac{1}{2}l'j'). \tag{23}$$

For the magnetic multipole we have the general matrix elements

$$\frac{1}{ik}e\langle M_k\rangle\cdot\left[\int\psi_{ljm}{}^*\boldsymbol{\sigma}\cdot(\mathbf{L}r^{-k-1}C^{(k)})\phi_{l'j'm'}dv+\int\phi_{ljm}{}^*\boldsymbol{\sigma}\cdot(\mathbf{L}r^{-k-1}C^{(k)})\psi_{l'j'm'}dv\right].$$

But we can write

$$\boldsymbol{\sigma}\cdot(\mathbf{L}r^{-k-1}C_\mu{}^{(k)})\phi=\boldsymbol{\sigma}\cdot\mathbf{L}(r^{-k-1}C_\mu{}^{(k)}\phi)-r^{-k-1}C_\mu{}^{(k)}(\boldsymbol{\sigma}\cdot\mathbf{L}\phi);$$

and also, by the Hermiticity of $\boldsymbol{\sigma}\cdot\mathbf{L}$,

$$\int\psi^*\boldsymbol{\sigma}\cdot\mathbf{L}(r^{-k-1}C_\mu{}^{(k)}\phi)dv=\int(\boldsymbol{\sigma}\cdot\mathbf{L}\psi)^*r^{-k-1}C_\mu{}^{(k)}\phi dv.$$

Now the functions ψ,ϕ are eigenfunctions of the operator $\boldsymbol{\sigma}\cdot\mathbf{L}$ belonging to the eigenvalues $\kappa-1$, $\bar{\kappa}-1$ respectively; where κ is the Dirac quantum number

$$\kappa=\begin{cases}j+\frac{1}{2} & j=l+\frac{1}{2}\\ -(j+\frac{1}{2}) & j=l-\frac{1}{2}\end{cases}$$

referred to the large component and $\bar{\kappa}=-\kappa$. We thus get, for the reduced matrix elements of the magnetic multipoles,

$$-(e/k)\langle\|M_k\|\rangle(\kappa+\kappa')\times\int_0^\infty r^{-k-1}(fg'+gf')dr(\tfrac{1}{2}lj\|C^{(k)}\|\tfrac{1}{2}\bar{l}'j'), \tag{24}$$

which have the parity selection ($l+l'+k$ odd). The general reduced matrix elements of the $C^{(k)}$ can be calculated best with the techniques of Racah.[2] The result is (using a formula of Schwinger[10])

$$(\tfrac{1}{2}lj\|C^{(k)}\|\tfrac{1}{2}l'j')=\tfrac{1}{2}[1+(-1)^{l+l'+k}]\Delta(jj'k)(-1)^{(j'-j+s-1)/2}\times\frac{(j+j'+s)!!}{(j+j'-s)!!(j-j'+s-1)!!(j'-j+s-1)!!}, \tag{25}$$

where

$$s=\begin{cases}k & j+j'+k \text{ even}\\ k+1 & j+j'+k \text{ odd}\end{cases}$$

and

$$\Delta(abc)=\left[\frac{(a+b-c)!(b+c-a)!(c+a-b)!}{(a+b+c+1)!}\right]^{\frac{1}{2}}.$$

For the first order hyperfine interactions only the diagonal matrix elements are needed and we get for

[10] J. Schwinger, *On Angular Momentum* (Nuclear Development Associates, Inc., White Plains, 1952), pp. 34, 35.

the interaction constants:

$$A_k=-Q_k e\int_0^\infty r^{-k-1}(f^2+g^2)dr(-1)^{k/2}\frac{(k-1)!!}{k!!}\times\frac{(2j)!}{(2j+k)!!(2j-k-1)!!},\quad k \text{ even}, >0; \quad (26a)$$

$$A_k=-M_k\frac{e}{k}4\kappa\int_0^\infty r^{-k-1}fgdr(-1)^{(k-1)/2}\frac{k!!}{(k-1)!!}\times\frac{(2j)!}{(2j+k+1)!!(2j-k)!!},\quad k \text{ odd}, \quad (26b)$$

in terms of the specific nuclear moments

$$Q_k=\langle Q_k^0\rangle_{II}\quad M_k=\langle M_k^0\rangle_{II}. \quad (27)$$

The preceding analysis was for a single electron bound to the nucleus. It is also correct to describe the interaction of a single valence electron outside closed sub-shells of other electrons.[11] For configurations such as p^5, d^9 in $L-S$ coupling, or $(3/2)^3$, $(5/2)^5$ in j-j coupling, where there is just one electron less than the number needed to fill a shell, only very slight modifications are needed to give the correct matrix elements: the even (electric multipole) interactions are just (-1) times the values for a single electron while the odd (magnetic multipole) ones are the same.

In the case of more complex electronic configurations one must know the coupling scheme of the several angular momenta involved; then the techniques of Racah[2] show how to calculate the appropriate "projection" factors.

RADIAL INTEGRALS

With the separation of variables (21) the Dirac equation (20) for the radial functions f and g reads

$$\left(\frac{d}{dr}-\frac{\kappa}{r}\right)f=\frac{1}{\hbar c}(mc^2+E+eV_c)g,$$
$$\left(\frac{d}{dr}+\frac{\kappa}{r}\right)g=\frac{1}{\hbar c}(mc^2-E-eV_c)f. \quad (28)$$

For a many-electron atom the best solution consistent with the assumption of the preceding footnote is obtained from a Hartree-Fock treatment. However, to obtain simple analytical results we make the assumption, following Casimir,[3] that the important contribution to the integrals (26) comes from the region of small values of r. This should be an excellent approximation for the cases $j=l+\frac{1}{2}=k/2$ (magnetic dipole in $s_{\frac{1}{2}}$ state, magnetic octupole in $p_{\frac{1}{2}}$ state, etc.) where the nonrelativistic treatment gives the interaction as due entirely to the electron density at the nucleus ($r=0$). For orbits of larger l, however, the wave function is concentrated farther out and is more slowly rising near the origin so this approximation worsens. At small value of r the major contribution to the potential is from the nucleus. Setting $V_c=Ze/r$ and with the approximation of zero binding energy $|mc^2-E|\ll V_c$, we get from (28) the solutions in terms of Bessel functions:

$$f=C[\tfrac{1}{2}xJ_{2\rho+1}(x)-(\rho+\kappa)J_{2\rho}(x)],$$
$$g=C\alpha ZJ_{2\rho}(x), \quad (29)$$

where $x=(8Zr/a_0)^{\frac{1}{2}}$, $\rho=(\kappa^2-\alpha^2Z^2)^{\frac{1}{2}}$, $a_0=\hbar^2/me^2\alpha=e^2/\hbar c$. With these functions the radial integrals (26) can be evaluated[12] to give the following results:

$$\int_0^\infty r^{-k-1}(f^2+g^2)dr=C^2\left(\frac{2Z}{a_0}\right)^k\frac{(2k-2)!}{k!(k-1)!}\times\frac{[k(2\kappa+k)(2\kappa+k-1)-4\alpha^2Z^2(3k-1)]}{(2\rho+k)(2\rho+k-1)\cdots(2\rho-k)}, \quad (30)$$

$$\int_0^\infty r^{-k-1}fgdr=C^2\frac{h}{2mc}\left(\frac{2Z}{a_0}\right)^{k+1}\frac{(2k-1)!}{k!(k-1)!}\times\frac{(-k-2\kappa)}{(2\rho+k)(2\rho+k-1)\cdots(2\rho-k)}. \quad (31)$$

The normalization constant C, which gives the density at the nucleus of the wave function of the outer valence electron, is best evaluated in terms of the fine-structure separation (for non-s electrons) between the states $j=l+\frac{1}{2}$ and $j=l-\frac{1}{2}$, which have almost identical wave functions for larger values of r. Here and subsequently we shall use the notation of a single ($'$) to identify a quantity as relating to the state $j=l+\frac{1}{2}$, and a double ($''$) for the state $j=l-\frac{1}{2}$. The resulting identification is[3-5]

$$C'\cong -C'',\quad C^2=\frac{\delta}{2.911}\frac{l(l+1)}{2HZ^3}\frac{1}{a_0}, \quad (32)$$

where δ is the doublet splitting in cm^{-1}, and H is a relativistic correction factor. More accurate approximations for the ratio $|C''/C'|$ will be termed normalization corrections and will be of concern in the following section. Casimir[3] gives the estimate

$$\left|\frac{C''}{C'}\right|^2=1+\frac{3\alpha^2Z^2}{2l(l+1)n^*} \quad (33)$$

[11] This assumes that one can write the total wave function for all the electrons in the form of products where the coordinates of the valence electron are separated from those of the core electrons.

[12] G. N. Watson, *Theory of Bessel Functions* (Cambridge University Press, London, 1952), p. 403.

involving the "effective quantum number" n^*. However, an explicit calculation by Breit for the case of thallium ($Z=81$) gives a value $|C''/C'|^2=1.65$ compared to the 1.18 of (33). We shall use Casimir's formula (33) for lighter nuclei ($Z\lesssim 50$) for which these corrections are not very large anyway. For the integrals of greatest interest we shall write the results in the following forms.

Magnetic dipole:

$$\int_0^\infty r^{-2}fgdr=C^2\frac{\hbar}{2mc}\left(\frac{2Z}{a_0}\right)^2\frac{F}{(2l+1)[\mp(2l+1)-1]},$$

$$\mp \text{ as } j=l\pm\tfrac{1}{2}. \quad (34a)$$

Electric quadrupole:

$$\int_0^\infty r^{-3}(f^2+g^2)dr=C^2\left(\frac{2Z}{a_0}\right)^2\frac{R}{l(2l+1)(2l+2)}. \quad (34b)$$

Magnetic octupole:

$$\int_0^\infty r^{-4}fgdr=C^2\frac{\hbar}{2mc}\left(\frac{2Z}{a_0}\right)^4 \times\frac{10T}{(2l+3)(2l+2)(2l+1)(2l)(2l-1)[\mp(2l+1)-3]},$$

$$\mp \text{ as } j=l\pm\tfrac{1}{2}. \quad (34c)$$

F and R are the same relativistic correction factors given by Casimir[3]; T is the corresponding correction factor for the octupole integral and is given by

$$T=\frac{(2j+4)!}{(2j-3)!}\frac{(2\rho-4)!}{(2\rho+3)!}. \quad (35)$$

All these factors, along with H, are plotted as functions of Z in Fig. 1 for the case $l=1$.

SECOND-ORDER EFFECTS

So long as we consider only first order effects of the hfs interactions, the multipoles can be separated from one another unambiguously by the orthogonality of the "interval rules" for different orders (8). However in second order we get the energy given by the square of a matrix element. Thus if, in second order, we consider the matrix element from the state IJF to the (different) state $I'J'F$ of the hfs interactions of various orders, we get a dependence on F which goes as the Racah coefficient,

$$(-1)^F W(I'J'IJ;Fk).$$

In the square of the matrix element there will be such products as

$$W(I'J'IJ;Fk_1)W(I'J'IJ;Fk_2),$$

and if we want to know what part of this looks like the first order term of an interaction of rank k, we multiply by

$$(-1)^F W(IJIJ;Fk)(2F+1),$$

and sum over F. This sum is well known in the theory of Racah coefficients and gives a result proportional to

$$W(Ik_1Ik_2;I'k)W(Jk_1Jk_2;J'k),$$

which is nonzero only if

$$|k_1-k_2|\leq k\leq k_1+k_2.$$

Thus in second order the square of the dipole term can influence at most the quadrupole; the cross dipole-quadrupole term can affect up to the octupole; and the square of the quadrupole term can reach to the 2^4-pole.

We shall now calculate the off-diagonal matrix elements for the dipole and quadrupole operators from the state in which the measurement is being made (assumed to be $j=l+\frac{1}{2}$) to the near-by doublet level of the electron ($j=l-\frac{1}{2}$).

For the dipole term the matrix element (always diagonal in F) is, from (24),

$$W(IJIJ-1;F1)(-1)^{I+J-F-1}\langle I\|M_1\|I\rangle(-e)(\kappa'+\kappa'')\times\int_0^\infty r^{-2}(f'g''+g'f'')dr(\tfrac{1}{2}lJ\|C^{(1)}\|\tfrac{1}{2}lJ-1)$$

and from (25),

$$(\tfrac{1}{2}lJ\|C^{(1)}\|\tfrac{1}{2}lJ-1)=[(2J+1)(2J-1)/4J]^{\frac{1}{2}};$$

also $\kappa'+\kappa''=1$.

The form of the Racah coefficient is

$$W(IJIJ-1;F1)(-1)^{I+J-F-1}=\left(\frac{(I+J-F)(J-I+F)(I-J+F+1)(I+J+F+1)}{(I+1)(2I+1)2I(2J+1)2J(2J-1)}\right)^{\frac{1}{2}}$$

and the nuclear term is

$$\langle I\|M_1\|I\rangle=[(I+1)(2I+1)/I]^{\frac{1}{2}}M_1,$$

so that the entire matrix element is

$$\frac{-eM_1}{4IJ}\int_0^\infty r^{-2}(f'g''+g'f'')dr[(I+J-F)(J-I+F)\times(I-J+F+1)(I+J+F+1)]^{\frac{1}{2}}. \quad (36)$$

The radial integral yields

$$\int_0^\infty r^{-2}(f'g''+g'f'')dr=C'C''\frac{\hbar}{2mc}\left(\frac{2Z}{a_0}\right)^2\times\frac{-4\Gamma(\rho'+\rho''-1)}{\Gamma(\rho''-\rho'+2)\Gamma(\rho'-\rho''+2)\Gamma(\rho'+\rho''+2)}\equiv-C'C''\frac{\hbar}{2mc}\left(\frac{2Z}{a_0}\right)^2\frac{G}{l(2l+1)(2l+2)}. \quad (37)$$

The ratio of this to the diagonal term in the state $j=l+\frac{1}{2}$ is

$$\int_0^\infty r^{-2}(f'g''+g'f'')dr \Big/ \int_0^\infty r^{-2}f'g'dr = \frac{C''}{C'}\frac{1}{l}\frac{G}{F'} \equiv -\frac{1}{l}\xi. \quad (38)$$

The off-diagonal quadrupole matrix element is, from (23),

$$W(IJIJ-1;F2)(-1)^{I+J-F-1}\langle I\|Q_2\|I\rangle(-e) \times \int_0^\infty r^{-3}(f'f''+g'g'')dr\,(\tfrac{1}{2}lJ\|C^{(2)}\|\tfrac{1}{2}lJ-1)$$

and, from (25),

$$(\tfrac{1}{2}lJ\|C^{(2)}\|\tfrac{1}{2}lJ-1)=\left(\frac{3(2J+1)(2J-1)}{16J(J+1)(J-1)}\right)^{\frac{1}{2}}.$$

The form of the Racah coefficient is

$$W(IJI\,J-1;F2)(-1)^{I+J-F-1} =[F(F+1)-I(I+1)-J^2+1]\left(\frac{3(I+J-F)(J-I+F)(I-J+F+1)(I+J+F+1)}{4J(J+1)(J-1)(2J-1)(2J+1)(2I+3)(2I+1)(2I-1)I(I+1)}\right)^{\frac{1}{2}},$$

and the nuclear term is

$$\langle I\|Q_2\|I\rangle=\left(\frac{(2I+3)(2I+1)(I+1)}{I(2I-1)}\right)^{\frac{1}{2}}Q_2,$$

so the entire matrix element is

$$-eQ_2\int_0^\infty r^{-3}(f'f''+g'g'')dr[(I+J-F)(J-I+F) \times(I-J+F+1)(I+J+F+1)]^{\frac{1}{2}} \times[F(F+1)-I(I+1)-J^2+1] \times\frac{3}{8J(J+1)(J-1)I(2I-1)}. \quad (39)$$

The radial integral gives

$$\int_0^\infty r^{-3}(f'f''+g'g'')dr =C'C''\left(\frac{2Z}{a_0}\right)^2\frac{2\Gamma(\rho'+\rho''-2)}{\Gamma(\rho'+\rho''+3)\Gamma(\rho'-\rho''+3)\Gamma(\rho''-\rho'+3)} \times\{12[\alpha^2Z^2+(\rho'+\kappa')(\rho''+\kappa'')]+(\rho'+\rho''-1) \times(\rho'+\rho''-2)(\rho'-\rho''+2)(\rho''-\rho'+2) -3(\rho'+\kappa')(\rho'+\rho''-2)(\rho'-\rho''+2) -3(\rho''+\kappa'')(\rho'+\rho''+2)(\rho''-\rho'+2)\} \equiv -C'C''\left(\frac{2Z}{a_0}\right)^2\frac{S}{l(2l+1)(2l+2)}. \quad (40)$$

The ratio of this to the diagonal integral in the $j=l+\frac{1}{2}$ state is

$$\int_0^\infty r^{-3}(f'f''+g'g'')dr \Big/ \int_0^\infty r^{-3}(f'^2+g'^2)dr = -\frac{C''}{C'}\frac{S}{R'}\equiv\eta. \quad (41)$$

Collecting all the terms, we can now write the second order energy as follows:

$$W_F^{(2)}=\frac{1}{\Delta E}(I+J-F)(J-I+F) \times(I-J+F+1)(I+J+F+1) \times\left[-\frac{3(F(F+1)-I(I+1)-J^2+1)}{2J(J-1)(2J-1)I(2I-1)}\eta A_2' -\frac{J+1}{2IJ(2J+1)(2J-1)}\xi A_1'\right]^2, \quad (42)$$

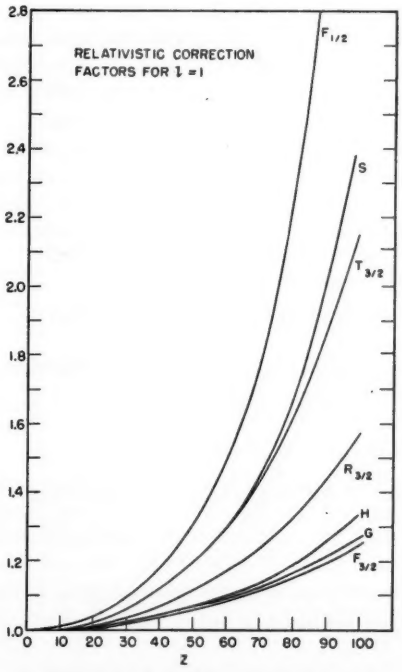


FIG. 1. Relativistic correction factors for $l=1$.

in terms of the first-order interaction constants in the state $j=l+\frac{1}{2}$ (we might also have referred to the state $j=l-\frac{1}{2}$). ΔE is $-\delta$ (the fine-structure splitting) if the $j=l+\frac{1}{2}$ state is the lower state (in energy), or $+\delta$ if $j=l-\frac{1}{2}$ is lower.

EFFECTS OF CONFIGURATION INTERACTION

We now go on to consider the effect of some configuration interaction of the sort discussed by Fermi and Segrè[13] and calculated in a particular case by Koster.[14] For configurations s^2lj (or $s^2l^{-1}j$) we include the possibility of one of the s electrons being raised to a higher s-state s'. The wave function in $L-S$ coupling will be written—for both $j=l+\frac{1}{2}$ and $j=l-\frac{1}{2}$ levels

$$\Psi_j=\alpha_0(s^2(S=0)^2L_j)+\alpha_1(ss'(S=1)^2L_j)+\alpha_2(ss'(S=0)^2L_j) \quad (43)$$

with normalization $\alpha_0^2+\alpha_1^2+\alpha_2^2=1$, where S is the resultant angular momentum of the two s electrons' spins which then couples to the spin of the l electron to give the doublet. In what follows we shall approximate only that $\alpha_1^2\ll 1$ (Koster finds $\alpha_1^2=0.001$ for gallium, $Z=31$).

For the wave function (43) the octupole and quadrupole matrix elements, as well as the fine-structure are essentially the same (to order α_1^2) as those one would get from considering only the valence l electron alone. We are interested in the effect of the s-electrons in the first and second order dipole interactions as these influence the interpretation of the purely octupole interaction from the hfs data. We shall find an explicit evaluation for a correction factor which should be multiplied into A_1' in formula (42) just to take account of the dipole interaction of these s electrons.

First, with the total dipole operator written as a sum of an operator $T_l^{(1)}$ (of rank 1) acting on the valence l electron and another $T_s^{(1)}$ acting on the s electrons, the general reduced matrix element becomes (to order $\alpha_1^2\ll 1$)

$$(J\|T_l^{(1)}+T_s^{(1)}\|J')=(J\|T_l^{(1)}\|J')+\Delta_{JJ'}, \quad (44)$$

where $\Delta_{JJ'}$ is a sum of matrix elements between various terms of (43), all of the form

$$\Delta_{JJ'}\sim(S\tfrac{1}{2},\tfrac{1}{2}l,J\|T_s^{(1)}\|S'\tfrac{1}{2},\tfrac{1}{2}l,J')$$
$$=W(\tfrac{1}{2}J\tfrac{1}{2}J';l1)(2J+1)^{\frac{1}{2}}(2J'+1)^{\frac{1}{2}}(-1)^{l-\frac{1}{2}+J}\times(S\tfrac{1}{2},\tfrac{1}{2}\|T_s^{(1)}\|S'\tfrac{1}{2},\tfrac{1}{2}).$$

That is, without actually calculating $\Delta_{JJ'}$ we have gotten its dependence on J and J'. Now, putting $J=l+\frac{1}{2}$, we get the ratios

$$\frac{\Delta_{JJ-1}}{\Delta_{JJ}}=\left(\frac{2J-1}{J+1}\right)^{\frac{1}{2}}, \quad (45a)$$

$$\frac{\Delta_{J-1\,J-1}}{\Delta_{JJ}}=-\left(\frac{(J-1)(2J-1)}{(J+1)(2J+1)}\right)^{\frac{1}{2}}. \quad (45b)$$

Also the ratio of the off-diagonal to diagonal ($J=l+\frac{1}{2}$ state) reduced dipole matrix elements of the l electron is

$$-\xi\frac{[(J+1)(2J-1)]^{\frac{1}{2}}}{(2J+1)(2J-1)}. \quad (46)$$

The desired correction factor ζ is given by

$$\zeta=\frac{(lJ\|T_l^{(1)}\|lJ')+\Delta_{JJ'}}{(lJ\|T_l^{(1)}\|lJ')}, \quad (47)$$

and from now on we will understand $J=l+\frac{1}{2}$, $J'=l-\frac{1}{2}$ $=J-1$.

One must calculate Δ_{JJ} by taking the discrepancy between the observed interaction constant A_1' and that amount calculated for the valence l electron alone. If the hfs is measured in the $J'=l-\frac{1}{2}$ state as well, one can get a better check on Δ by solving the two simultaneous equations of the form (44) with the measured interaction constants A_1' and A_1''. Using (45b) and the relation

$$(lJ-1\|T_l^{(1)}\|lJ-1)=\left(\frac{(J+1)(2J-1)}{(J-1)(2J+1)}\right)^{\frac{1}{2}}\theta(lJ\|T_l^{(1)}\|lJ) \quad (48)$$

$$\theta=\frac{F''}{F'}\left|\frac{C''}{C'}\right|^2,$$

we get

$$M_1\Delta_{JJ}=\frac{\theta A_1'-[J/(J+1)]A_1''}{\theta+[(J-1)/(J+1)]}\times\left(\frac{(J+1)(2J+1)}{J}\right)^{\frac{1}{2}}, \quad (49a)$$

$$M_1(lJ\|T_l^{(1)}\|lJ)=\frac{[(J-1)/J]A_1'+A_1''}{\theta+[(J-1)/(J+1)]}\times\left(\frac{J(2J+1)}{(J+1)}\right)^{\frac{1}{2}}, \quad (49b)$$

and, finally,

$$\zeta=1+\frac{1}{\xi}\frac{[J/(J+1)]A_1''-\theta A_1'}{A_1''+[(J-1)/J]A_1'}\frac{(2J+1)(2J-1)}{J}. \quad (49c)$$

The calculations carried out here also find application in the study of the Zeeman effect in hfs as used to

[13] E. Fermi and E. Segrè, Z. Physik 82, 729 (1933).
[14] G. F. Koster, Phys. Rev. 86, 148 (1952).

measure directly the nuclear g factor. When an atom of spin $J=\frac{1}{2}$ (for $I>\frac{1}{2}$) is placed in a uniform magnetic field H, there are according to the Breit-Rabi formula pairs of lines arising from the hfs, the difference of whose frequencies gives directly the quantity $2g_I\mu_N H$.

Foley[15] has shown that for a $p_{\frac{1}{2}}$ electron state second-order contributions involving the doublet $p_{\frac{3}{2}}$ level can change the apparent value of g_I—as compared with the value measured directly by nuclear resonance methods. His formula is

$$R^{-1}=\frac{g_I(\text{atomic beam=hfs})}{g_I(\text{nuclear resonance})}=1-\frac{1837\Delta\nu}{6(2I+1)g_I\delta}, \quad (50)$$

where $\Delta\nu$ is the hfs interval in the $p_{\frac{1}{2}}$ state at zero field and δ is the fine-structure separation. Clendenin[16] has done the calculation relativistically and he gets formula (50) with the factor G/F'' included in the second term.

What enters in (50) is just the off-diagonal matrix element of the hfs interactions between the $p_{\frac{1}{2}}$ and $p_{\frac{3}{2}}$ states times the matrix element of the electron's magnetic moment operator between the same two states. There are three effects not considered by these other authors which we can now include: the normalization correction factor; the off-diagonal quadrupole term; the effect of configuration interaction on the off-diagonal dipole term. Using (42) we get the result

$$R^{-1}=1-\frac{1837}{g_I\delta}\left[\frac{\Delta\nu}{6(2I+1)}\zeta\frac{\xi}{\theta}-\frac{b}{I(2I-1)}\eta\right]. \quad (51)$$

The quantity b is the usual quadrupole interaction constant ($b=4A_2$) measured in the $p_{\frac{3}{2}}$ state and all other factors in (51) are as earlier defined. The sign of the correction term above is correct only when the $p_{\frac{1}{2}}$ state is lower in energy than the $p_{\frac{3}{2}}$ state.

We shall compare the calculated and measured values of this discrepancy for the ground states of gallium and indium.

Gallium[17]:

$$Z=31,\quad \delta=24.8\,10^6\text{ Mc/sec},\quad n^*=1.51,$$
$$|C''/C'|^2=1.02_5,\quad \xi=1.02,\quad \eta=1.04,\quad \theta=1.10_5,$$
$$\zeta=1.58,\quad A_1''/A_1'=2.34.$$

Ga^{69}:

$$\Delta\nu=2677\text{ Mc/sec},\quad b=62.5\text{ Mc/sec},\quad g_I=1.34,$$
$$R^{-1}=1-0.0078,$$

to be compared with the experimental value

$$1-0.0079\pm0.0023.$$

Ga^{71}:

$$\Delta\nu=3402\text{ Mc/sec},\quad b=39.4\text{ Mc/sec},\quad g_I=1.70,$$
$$R^{-1}=1-0.0084,$$

with the experimental value

$$1-0.0077\pm0.0017.$$

Indium[18]:

$$Z=49,\quad \delta=66.5\,10^6\text{ Mc/sec},\quad n^*=1.53,$$
$$|C''/C'|^2=1.06,\quad \xi=1.04,\quad \eta=1.11,\quad \theta=1.30,$$
$$\zeta=1.84,\quad A_1''/A_1=3.12.$$

In^{115}:

$$\Delta\nu=11\,330\text{ Mc/sec},\quad b=450\text{ Mc/sec},\quad g_I=1.22,$$
$$R^{-1}=1-0.0060,$$

and the experimental value is

$$1-0.0062\pm0.0005.$$

NUCLEAR MOMENTS

The nuclear moments are defined as the following expectation values (evaluated in the state $m_I=I$).

$$Q_k=\left(eg r^k C^{(k)}(\theta,\varphi)\right)_{II} \quad (52a)$$

for electric moments (k even);

$$M_k=\left(\mu_N(\nabla r^k C^{(k)}(\theta,\varphi))\cdot\left(g_l\frac{2}{k+1}\mathbf{L}+g_s\mathbf{S}\right)\right)_{II} \quad (52b)$$

for magnetic moments (k odd).

The magnetic multipole moments (52b) can also be written in the form

$$M_k=-\int r^k C^{(k)}(\theta,\varphi)\,\text{div}\mathbf{M}dv,$$

where $\mathbf{M}$ is the magnetization density (in the state $m_I=I$) defined as in Blatt and Weisskopf,[7] Chap. I. These are related to the usually defined moments as follows:

$\mu=M_1$, magnetic dipole moment;

$Q=2Q_2$, electric quadrupole moment;

and we shall define the magnetic octupole moment Ω as

$$\Omega=-M_3.$$

It can be seen from the phase factors in Eqs. (26a, b) that the moments of a given type, electric or magnetic, have a natural oscillation in sign as one proceeds to higher orders. The minus sign is introduced in the definition of Ω so that a nucleus with a positive dipole moment is most likely to have a positive octupole moment as well.

[15] H. M. Foley, Phys. Rev. **80**, 288 (1950).
[16] W. W. Clendenin, Phys. Rev. **94**, 1590 (1954).
[17] Data from G. E. Becker and P. Kusch, Phys. Rev. **73**, 584 (1948), and reference 15.
[18] Data from sources quoted in reference 15 and others given by P. Kusch in private communication.

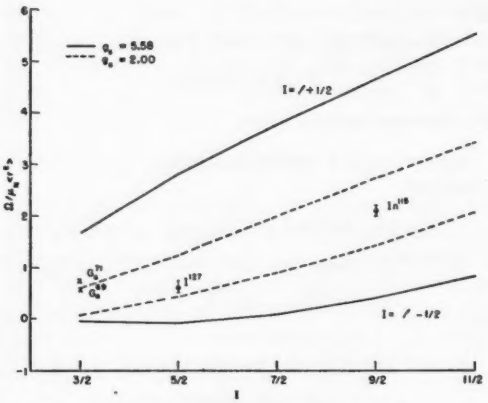


FIG. 2. Magnetic octupole moments of odd-proton nuclei.

It is of interest to calculate the moments expected of a single odd nucleon in an orbit of spin I. From (25) and (4b), we get directly the electric moments

$$Q_2=\frac{1}{2}Q=-\frac{1}{2}\frac{2I-1}{2I+2}eg_l\langle r^2\rangle, \tag{53a}$$

$$Q_4=-\frac{3}{8}\frac{(2I-1)(2I-3)}{(2I+4)(2I+2)}eg_l\langle r^4\rangle. \tag{53b}$$

For a nuclear configuration of n (odd) equivalent nucleons in the expected ground state, we have the relation

$$Q_k(j^nI=j)=\frac{2j+1-2n}{2j-1}Q_k(j), \tag{53c}$$

giving the moment of the several particles in terms of the value for a single particle.

The calculation of the magnetic multipole expectation values (52b) is slightly more involved. With extensive use of the Racah techniques we have derived the following general formula for matrix elements of this type in single-particle orbits; g is any function of r.

$$\begin{aligned}(l\tfrac{1}{2}j\|(\nabla gC^{(k)})\cdot(g_l\mathbf{L}+g_s\mathbf{S})\|l'\tfrac{1}{2}j')&\\
=\tfrac{1}{2}(1-(-1)^{l+l'+k})\{g_l\tfrac{1}{2}g/r(-1)^{j'-\frac{1}{2}+k}[k(k+1)&\\
-((j+\tfrac{1}{2})+(-1)^{j+j'+k}(j'+\tfrac{1}{2}))((j+\tfrac{1}{2})&\\
+(-1)^{j+j'+k}(j'+\tfrac{1}{2})+(-1)^{l+\frac{1}{2}+j})]&\\
+g_s\tfrac{1}{2}(-1)^{l'+j+j'}[(-1)^{l+\frac{1}{2}+j}dg/dr-(g/r)((j+\tfrac{1}{2})&\\
+(-1)^{j+j'+k}(j'+\tfrac{1}{2}))]\}\Delta(jj'k)(-1)^{(j'+j+s)/2+1}&\\
\times\frac{(j+j'+s)!!}{(j+j'-s)!!(j-j'+s-1)!!(j'-j+s-1)!!},&\end{aligned} \tag{54}$$

with s and $\Delta(\ )$ defined as in (25).

For the diagonal matrix elements in the state $M_I=I$, using (4b) and (54) we get

$$\mu=M_I=\mu_N I\begin{cases}g_l+(g_s-g_l)/2I, & I=l+\frac{1}{2};\\ g_l-(g_s-g_l)/(2I+2), & I=l-\frac{1}{2}.\end{cases} \tag{55}$$

the usual Schmidt values; and for the octupole

$$\Omega=-M_3=+\mu_N\frac{3}{2}\frac{(2I-1)}{(2I+4)(2I+2)}\langle r^2\rangle \times\begin{cases}(I+2)[(I-\frac{3}{2})g_l+g_s], & I=l+\frac{1}{2};\\ (I-1)[(I+5/2)g_l-g_s], & I=l-\frac{1}{2}.\end{cases} \tag{56}$$

One can make a plot of these values of the single-particle octupole moments very much like the Schmidt plot for dipoles. In Fig. 2 are the lines for $I=l+\frac{1}{2}$ and $I=l-\frac{1}{2}$ of the quantity

$$\frac{\Omega}{\mu_N\langle r^2\rangle}\text{ as a function of }I(\geqslant\tfrac{3}{2})$$

for an odd proton ($g_l=+1$, $g_s=5.58$); a similar plot can be drawn for an odd neutron ($g_l=0$, $g_s=-3.83$).

For nuclear configurations of several equivalent particles in the ground state $I=j$ we get for the magnetic multipole moments

$$M_k(j^nI=J)=M_k(j). \tag{57}$$

If (as in an odd-odd nucleus for example) we have a configuration of two particles (or two separate groups of particles) with separate spins j_1 and j_2 coupled to a resultant I, any multipole moment of the total system is made up out of the moments of the two particles as follows:

$$\begin{aligned}(j_1j_2Im_I=I|T_1^{(k)}+T_2^{(k)}|j_1j_2Im_I=I)&\\
=\frac{(2I+1)!}{[(2I-k)!(2I+k+1)!]^{\frac{1}{2}}}\Bigg\{\frac{[(2j_1-k)!(2j_1+k+1)!]^{\frac{1}{2}}}{(2j_1)!}&\\
\times(j_1m_1=j_1|T_1^{(k)}|j_1m_1=j_1)&\\
\times W(j_1Ij_1I;j_2k)(-1)^{j_2-j_1-I+k}&\\
+\frac{[(2j_2-k)!(2j_2+k+1)!]^{\frac{1}{2}}}{(2j_2)!}&\\
\times(j_2m_2=j_2|T_2^{(k)}|j_2m_2=j_2)&\\
\times W(j_2Ij_2I;j_1k)(-1)^{j_1-j_2-I+k}\Bigg\}.&\end{aligned} \tag{58}$$

We can make one interesting and simple remark concerning the interpretation of nuclear moments in A. Bohr's asymmetric core model. In the strong-coupling situation where the valence nucleons are aligned with respect to a permanently distorted nuclear core we must reduce all the moments by a projection factor P_k which allows for the transformation of the necessary

operators into the body frame of the core. This projection factor, in the nuclear ground state where the valence nucleon is aligned with the core axis, is given by

$$P_k=(2I+1)\frac{[(2I)!]^2}{(2I-k)!(2I+k+1)!}. \tag{59}$$

It is interesting that the higher P_k's can be quite small numbers if I is small ($I\sim k/2$). For the octupole, for example,

$$P_3=1/35,\quad I=3/2,$$
$$=5/42,\quad I=5/2,$$

while the smallest P_1 is $\frac{1}{3}$ (for $I=\frac{1}{2}$) and the smallest P_2 is $\frac{1}{5}$ (for $I=\frac{3}{2}$).

The contrast between the asymmetric core effects in quadrupole and octupole moments is further intensified by the fact that while it is the large numerical charge of the core which, in spite of the projection factor, creates a large quadrupole moment, the total magnetization of the core is only of the order of that produced by a single particle. The conclusion is that if the strong-coupling situation exists for nuclei with small spins (3/2,5/2) the octupole moment should be much smaller than the expected single particle value.

EXAMPLES: $p_{3/2}$ ELECTRON

For an electron state with a single valence electron in a $p_{\frac{3}{2}}$ orbit, there will be (for $I\geqslant\frac{3}{2}$) four hfs levels with the $M(IJ;Fk)$ coefficients (6), (7) given in Table I.

If we let x, y, z be the measured intervals between the $F=I+\frac{3}{2}-F=I+\frac{1}{2}$; $F=I+\frac{1}{2}-F=I-\frac{1}{2}$; $F=I-\frac{1}{2}-F=I-\frac{3}{2}$ levels respectively, then we get, for the interaction constants (8), $W_{I-\frac{3}{2}}=0$; $W_{I-\frac{1}{2}}=z$; $W_{I+\frac{1}{2}}=(y+z)$; $W_{I+\frac{3}{2}}=(x+y+z)$

$$A_1=\left[\frac{9}{20}\frac{I}{I+1}x+\frac{3}{20}\frac{(2I-1)(2I+3)}{(I+1)(I+2)}y+\frac{9}{20}\frac{(I-1)}{(I+2)}z\right]\frac{(2I+4)}{(2I+1)}, \tag{60a}$$

TABLE I. $M(IJ;Fk)$ coefficients.

	$k=1$	$k=2$	$k=3$
$F=I+\frac{3}{2}$	1	1	1
$F=I+\frac{1}{2}$	$\frac{I-3}{3I}$	$-\frac{(2I-1)(I+3)}{I(2I-1)}$	$-3\frac{2I+4}{2I}$
$F=I-\frac{1}{2}$	$\frac{-I-4}{3I}$	$-\frac{(2I+3)(I-2)}{I(2I-1)}$	$+3\frac{(2I+4)(2I+3)}{(2I)(2I-1)}$
$F=I-\frac{3}{2}$	$\frac{-3I-3}{3I}$	$\frac{(2I+3)(I+1)}{I(2I-1)}$	$-\frac{(2I+4)(2I+3)(2I+2)}{(2I)(2I-1)(2I-2)}$

$$A_2=-\frac{1}{2}\frac{I(2I-1)(I+2)}{(2I+1)(I+1)(2I+3)}x-\frac{1}{2}\frac{(2I-1)}{(2I+1)(I+1)}y-\frac{1}{2}\frac{(I-1)}{(2I+1)}z, \tag{60b}$$

$$A_3=\frac{1}{10}\frac{I(I-1)(2I-1)}{(2I+3)(2I+1)(I+1)}x-\frac{1}{10}\frac{(I-1)(2I-1)}{(2I+1)(I+1)}y+\frac{1}{10}\frac{(I-1)}{(2I+1)}z. \tag{60c}$$

We should subtract from the above formula for A_3 the amount due to the second order corrections (42); this comes to [using (8) to find the octupole-like part]:

$$\frac{1}{\Delta E}\frac{3}{10}\frac{I-1}{I}\eta A_2'\left[\frac{8}{2I-1}\eta A_2'-\frac{5}{6}\xi\zeta A_1'\right], \tag{61}$$

with

$$\zeta=1+\frac{16}{\xi}\frac{\frac{3}{5}A_1''/A_1'-\theta}{3A_1''/A_1'+1}. \tag{62}$$

The formula for A_3 in terms of the octupole moment is

$$A_3=\mu_0\Omega\frac{16}{5\cdot7\cdot9}\frac{T}{H}\frac{Z}{a_0^5}\frac{\delta}{2.911}, \tag{63}$$

or

$$\Omega=A_3\frac{H}{T}\frac{3.36\times10^{-17}}{\delta Z},$$

in units of Ω—nuclear magneton cm², A_3—Mc/sec, and δ—cm^{-1}.

For the ground state of iodine[1]: $n^*=1.14$

I^{127}:

$$Z=53,\quad \delta=7600\text{ cm}^{-1},\quad A_1'=3100\text{ Mc/sec},$$

$$A_2'=286.6\text{ Mc/sec},\quad |C''/C'|^2=1.10,\quad \xi=1.05,\quad \eta=1.13.$$

No measurements have been made on the $p_{\frac{1}{2}}$ state but it is expected that there will be considerably less configuration interaction in the halogens than in the corresponding Group III elements due to the tighter binding of the s-electrons. We will thus assume $\zeta=1$. The formula for A_3 with corrections is

$$A_3=\frac{1}{5\cdot7\cdot16}[5x-16y+14z]-0.00053\text{ Mc/sec}$$
$$=(0.00287\pm0.00037-0.00053)\text{ Mc/sec},$$

where we have taken the square root of the sum of the squares of the experimental errors in x, y, z (weighted as above) as the total error. Using (63) with $H=1.07$, $T=1.22$, we get $\Omega_{127}=(0.17\pm0.03)\times10^{-24}$ nuclear magneton cm².

With the value for the radial integral taken roughly as

$$\langle r^2\rangle=\tfrac{3}{5}R_0^2=\tfrac{3}{5}(0.135A^{\frac{1}{3}})^2\times10^{-24}\text{ cm}^2,$$

we get the value (0.62 ± 0.10) on the octupole diagram ($\sim\frac{1}{4}$ the expected single-particle value).

For the metastable $p_{\frac{1}{2}}$ state of indium 115, Kusch[19] has remeasured the intervals with extreme accuracy. Using the correction factors already worked out, we get for A_3:

$$A_3=\frac{7}{2200}[6x-16y+11z]+0.00109 \text{ Mc/sec}$$
$$=(0.000011\pm0.000032+0.00109) \text{ Mc/sec}.$$

With $H=1.065$ and $T=1.19$, the octupole moment is

$$\Omega_{115}=(0.31\pm0.01)\times10^{-24} \text{ nuclear magneton cm}^2.$$

Approximating as above for $\langle r^2\rangle$, we get the value (2.1 ± 0.1) on the octupole plot ($\sim\frac{1}{2}$ the single-particle value).

Daly[20] has remeasured the hfs of the $p_{\frac{1}{2}}$ state for the two stable isotopes of gallium. The several correction factors have already been quoted; we have

$$Ga^{69}: A_3=\frac{1}{400}[x-4y+5z]+0.0000336 \text{ Mc/sec}$$
$$=(50.2\pm3+33.6)\times10^{-6} \text{ Mc/sec},$$

$$Ga^{71}: A_3=\frac{1}{400}[x-4y+5z]+0.0000285$$
$$=(85.8\pm3+28.5)\times10^{-6} \text{ Mc/sec};$$

with $H=1.02_6$, $T=1.06_5$ we get the octupole moments

$$\Omega_{69}=(0.107\pm0.004)\times10^{-24} \text{ nuclear magneton cm}^2.$$
$$\Omega_{71}=(0.146\pm0.004)\times10^{-24} \text{ nuclear magneton cm}^2.$$

Estimating $\langle r^2\rangle$ as before, we get the values (0.58) for Ga^{69} and (0.77) for Ga^{71} on the octupole plot.

The values of the quantity $\Omega/\mu_N\langle r^2\rangle$ for these four nuclides are displayed in Fig. 2, and it is striking to see the similarity between the distribution of points on this diagram and that on the Schmidt plot for dipole moments. Any strong conclusions about the quantitative aspects of this comparison may as yet be unjustified, since the rough estimate

$$\langle r^2\rangle=\tfrac{3}{5}R_0^2, \quad R_0=1.35A^{\frac{1}{3}}\times10^{-13} \text{ cm}$$

should be replaced by the analytical evaluations of some reasonable shell model. However, it is interesting to compare the sizes of the octupole moments for the isotopic pair $Ga^{69,71}$. The heavier nucleus has larger dipole and octupole moments and smaller quadrupole moment, thus is consistently closer to the pure single-particle picture.

The author would like to thank Professor P. Kusch, Doctors V. Jaccarino, J. G. King and R. T. Daly for making their experimental data known to him before publication. It is also with pleasure that the advice and encouragement received by him from Professors V. F. Weisskopf and S. D. Drell are acknowledged. Much of the author's familiarity with the problems and techniques of the study of hyperfine structure has come from numerous discussions with Doctor Vincent Jaccarino and other members of Professor Zacharias' Atomic Beam Laboratory.

[19] P. Kusch, Phys. Rev. **94**, 1799 (1954).
[20] R. T. Daly, Jr., and J. H. Holloway, Phys. Rev. **96**, 539 (1954).

APPENDIX I. DISCUSSION OF APPROXIMATIONS

In this section we shall discuss several approximations made in the theoretical analysis of this paper in order to arrive at an estimate of the accuracy of the terms calculated.

A: The assumption that a many-electron atom can be described as a core of closed shells plus a few valence electrons is the essential starting point for any study of atomic multiplet structure, fine structure and hyperfine structure. The corrections to this model, termed configuration interaction, include the admixture of excited states for the core electrons, brought about through the electrostatic interactions among all the electrons. The calculations of Sternheimer[21] have attempted to account for these effects in the dipole and quadrupole hyperfine interactions, the magnitude of his correction factors being of the order of 10 percent. Notwithstanding the difficulties of the labor involved, a calculation, similar to Sternheimer's, for the octupole interaction would be valuable.

B: In the evaluation of the radial integrals the use of unshielded coulomb wave functions is an excellent approximation for the octupole integral in a $p_{\frac{1}{2}}$ state; but for the dipole and quadrupole integrals of $\langle r^{-3}\rangle$ there may be a sizeable error, especially in the lighter elements. As an example, integrating $\langle r^{-3}\rangle$ with a Hartree wave function[22] for gallium from $r=0$ to $r=0.05a_0$, one has only 50 percent of the entire $\langle r^{-3}\rangle$ integral while the strength of the central potential is already shielded by 20 percent. In calculating the second-order corrections to the hyperfine structure, only ratios of these $\langle r^{-3}\rangle$ integrals are needed, so the major part of this error is eliminated. For the best evaluation of these terms one might take values for ξ and η somewhere between unity and the values given in the text.

The uncertainty in the value of the normalization constant C^2 is not easy to evaluate. It would be interesting to check formulas (32) by carrying out the numerical solution of the Dirac radial equations with some reasonable approximation for the complex central field in several atoms.

The discussions A and B relate to the problem of getting the nuclear octupole moment from the corrected interaction constant, and as a figure of merit for the results used in the preceding section we suggest a value of about 15 percent.

[21] R. Sternheimer, Phys. Rev. **84**, 244 (1951).
[22] Hartree, Hartree, and Manning, Phys. Rev. **59**, 299 (1941).

C: The accuracy of the second order calculation involving the doublet state should be very good. The error is probably no more than a couple of percent for the terms relating to the valence l-electron (see B above) and very likely no more for the s-electron correction factor, all these quantities being derived from other experimental numbers with only slight theoretical correction. The only check on these several terms is in the explanation of the g_I-discrepancy in a $p_{\frac{1}{2}}$ state, where at present the large experimental uncertainties prevent a closer verification.

D: The biggest question in evaluating the second order corrections is about the contributions of other electronic levels besides the doublet state. One would like to rely on the larger energy denominators, ΔE_n, associated with all other terms of the perturbation sum to keep their contributions smaller by a factor $\delta/\Delta E_n$ than the contribution of the doublet level alone, but the total effect of the infinity of terms is not easily seen.

First, one can simplify the problem just a little with the following results. One can show in general that the octupole-like part of the (quadrupole)^2 term from a general $^2P_{\frac{1}{2}}$ state (in $L-S$ coupling) to any other perturbing ^2L_J state is zero if one adds the contributions of both doublet states $J=L+\frac{1}{2}$ and $J=L-\frac{1}{2}$. The only residual contribution of such terms would be due to the slightly different energy denominators of the two doublet states, thus an order of magnitude smaller than any straightforward estimate.

The (quadrupole)^2 term is anyway smaller than the cross dipole-quadrupole term and it is the latter one that we must worry about now. One might think that a useful estimate of this problem could be gotten from a closure approximation. That is, one tries to represent the second order sum as follows:

$$\sum_n \frac{|\langle i|H_1|n\rangle|^2}{\Delta E_n}=\frac{1}{\Delta E_{Av}}\sum_n|\langle i|H_1|n\rangle|^2, \qquad \text{(A1)}$$

where i refers to the initial state, n the intermediate states being summed over, and ΔE_{Av} is an average excitation energy for the particular problem.

For our problem, letting D and Q stand for the dipole and quadrupole operators, the second factor on the right-hand side of (A1) becomes the matrix element $\langle i|DQ|i\rangle$. The form of this operator is very much like the form of the octupole operator except that the product DQ has an extra factor e^2/r, which after taking the expectation value becomes a factor Ze^2/a_0. An upper limit for the evaluation of (A1) is gotten by setting $\Delta E_{Av}=\Delta E_{\min}\sim e^2/a_0$, which gives a result larger by a factor Z than the first order octupole matrix element.

It must be pointed out that equating ΔE_{Av} to $\Delta E_{\min}$ is an extremely bad approximation for our problem. The reason for this is that our operators are very strongly varying functions (r^{-3}) so that the correct average excitation energy ΔE_{Av} is some very high energy. By way of justifying this last statement we cite the example of a delta-function perturbation which requires an infinite value of ΔE_{Av} to make (A1) meaningful. We thus believe that the closure approximation is useless in our problem.

We will now try to carry out part of the second order sum in an approximate manner. First, the matrix element from a p-state to an f-state are exceedingly small compared with the p-p matrix elements. The octupole part of the dipole-quadrupole matrix product from a $p_{\frac{1}{2}}$ state to a $p_{\frac{1}{2}}$ state is, from (61),

$$-\frac{1}{4}\frac{I-1}{I}A_2A_1,$$

and the corresponding contribution from the $p_{\frac{1}{2}}-p_{\frac{1}{2}}$ matrix elements turns out to be

$$+\frac{1}{5}\frac{I-1}{I}A_2A_1, \qquad \text{(A2)}$$

where all the finer correction factors have been ignored. If we consider the two doublet levels of any perturbing ^2P state to have the same energy denominators, these two terms cancel strongly, leaving only $\frac{1}{5}$ of the original $p_{\frac{1}{2}}-p_{\frac{1}{2}}$ term.

We must also take into account the poorer overlap of the radial wave-functions as we proceed to higher perturbing levels. For bound states of a single valence electron Casimir gives the normalization constant C^2 for any level as proportional to n^{*-3}, where n^* is the effective quantum number for that level. Comparing the sum over all p-doublets up to zero energy with the value found in the ground state doublet alone, we have to evaluate

$$\sum_{i=1}^{\infty}\left(\frac{n^*}{n^*+i}\right)^3,$$

where n^* here refers to the ground state. This number is about 0.4 for $n^*=1.5$ and 0.2 for $n^*=1$. Combining these several factors we may estimate the value of the apparent octupole interaction due to all levels for the single electron up to $E=0$ as

$$\tfrac{1}{5}(0.4)\delta/\Delta E_{\min}$$

times the correction obtained from the ground state doublet alone. Values of $\Delta E_{\min}$ for several atoms are 1/300 for Al, 1/95 for Cl, 1/40 for Ga, 1/20 for Br, 1/13 for In, 1/9 for I, which result in corrections of less than one percent for all these atoms.

In summary, the discussions C and D relating to the accuracy of the second order corrections to the octupole interactions are still quite crude and incomplete. However in view of the optimistic results which these discussions do suggest, we will guess an accuracy of

about 5 percent for the corrections as calculated in the preceding section.

APPENDIX II; SKETCH OF THE NONRELATIVISTIC THEORY

For a nonrelativistic study, the hyperfine interactions may be conveniently described directly in terms of the two charge-current densities, without using the intermediary fields. Thus for the electric interaction we write the energy (to first order):

$$W_e=\int\int\frac{\rho_1\rho_2}{r_{12}}dv_1dv_2,$$

and with the assumption that system 1 is outside system 2 we get the multipole expansion

$$W_e=\sum_k\int\rho_1 r^{-k-1}C^{(k)}dv_1\cdot\int\rho_2 r^kC^{(k)}dv_2. \quad \text{(A3)}$$

Then identifying ρ as e times the wave function product $\psi^*\psi$ (A3) can be read as the product of two matrix elements.

For the magnetic interaction between two current systems, the interaction is

$$W_m=\frac{-1}{c^2}\int\int\frac{\mathbf{j}_1\cdot\mathbf{j}_2}{r_{12}}dv_1dv_2. \quad \text{(A4)}$$

However, because of the vector nature of the currents $\mathbf{j}$ we cannot immediately make a multipole expansion of this expression, (A4). We first express each current density $\mathbf{j}$ in terms of a magnetization density $\mathbf{M}$:

$$\mathbf{j}=c\ \mathrm{curl}\mathbf{M}.$$

Then a series of partial integrations reduces (A4) to

$$W_m=\int\int\frac{\mathrm{div}\mathbf{M}_1\ \mathrm{div}\mathbf{M}_2}{r_{12}}dv_1dv_2 \quad \text{(A5)}$$

provided the two systems 1 and 2 do not overlap anywhere. Now we can make the usual expansion to get

$$W_m=\sum_k\int\mathrm{div}\mathbf{M}_1 r^{-k-1}C^{(k)}dv_1\cdot\int\mathrm{div}\mathbf{M}_2 r^kC^{(k)}dv_2. \quad \text{(A6)}$$

The analysis of the angular dependence of the hfs interactions is just as before and we get for the interaction constants

$$A_k=e^2\langle r^{-k-1}C^{(k)}g_l\rangle_{JJ}\langle r^kC^{(k)}g_l\rangle_{II} \quad \text{(A7)}$$

for electric multipole, k even;

$$A_k=\langle r^{-k-1}C^{(k)}\ \mathrm{div}M_1\rangle_{JJ}\langle r^kC^{(k)}\ \mathrm{div}M_2\rangle_{II}$$

for magnetic multipole, k odd.

The magnetic terms are not yet in the desired form, ψ^* operator ψ. Using vector identities and carrying out some partial integrations (see reference 6), one can re-express the magnetization in terms of convection and spin currents through the operators $\mathbf{L}$ and $\mathbf{S}$. The final result for the magnetic multipoles is

$$A_k=\mu_0\mu_N\left\langle(\nabla r^{-k-1}C^{(k)})\cdot\left(\frac{-2}{k}g_l\mathbf{L}+g_s\mathbf{S}\right)\right\rangle_{JJ}$$
$$\times\left\langle(\nabla r^kC^{(k)})\cdot\left(\frac{2}{k+1}g_l\mathbf{L}+g_s\mathbf{S}\right)\right\rangle_{II}. \quad \text{(A8)}$$

For single electron states $\frac{1}{2}lJ$, the matrix elements occurring in (A7) and (A8) can be evaluated by using formulas (25) and (54) respectively. For the first four orders the results are ($g_l=-1$, $g_s=-2$ for electron):

$$A_1=IJa=\mu_0\frac{2l(l+1)}{J+1}\langle r^{-3}\rangle M_1, \quad \text{(A9)}$$

$$A_2=1/4b=e^2\frac{(2J-1)}{2(2J+2)}\langle r^{-3}\rangle Q_2, \quad \text{(A10)}$$

$$A_3=c=-\mu_0\frac{8l(l-1)(l+1)(l+2)}{(2J+2)(2J+3)(2J+4)}\langle r^{-5}\rangle M_3, \quad \text{(A11)}$$

$$A_4=d=-e^2\frac{3}{8}\frac{(2J-1)(2J-3)}{(2J+2)(2J+4)}\langle r^{-5}\rangle Q_4, \quad \text{(A12)}$$

where we have used the nuclear moments as defined in (18).

These formulas are invalid for the special case of magnetic 2^k pole interaction in an electron state $J=l+\frac{1}{2}=k/2$—dipole in $s_{\frac{1}{2}}$ state, octupole in $p_{\frac{3}{2}}$ state. For these cases an alternative analysis is carried out as follows. The vector potential,

$$\mathbf{A}(2)=\frac{1}{c}\int\frac{\mathbf{j}(1)}{r_{12}}dv_1,$$

is easily evaluated for the electron in the state $M_J=J$ considering the spin and convection current contributions to $\mathbf{j}$ in the usual way. Taking just the $k(=2l+1)$-pole term we find that the magnetic field which it represents can easily be written as the gradient of a scalar. That is

$$\mathbf{H}=\mathrm{curl}\mathbf{A}=\mathrm{grad}\varphi,$$

$$\varphi=-\mu_0g(0)(-1)^l\frac{2(2l)!(2l+1)!(2l+2)!}{l!l!(4l+3)!}r^kC^{(k)}, \quad \text{(A13)}$$

where, if $f(r)$ is the normalized radial wave function,

$g(r)=|r^{-1}f(r)|^2$. Now a formula for the magnetic interaction equivalent to (A4) is

$$W_m=-\int \mathbf{H}\cdot\mathbf{M}dv,$$

where $\mathbf{M}$ is the nuclear magnetization density. Putting in (A13) and performing one partial integration, we have the effective evaluation of the electronic matrix element (A7) for these special cases. Thus, for an $s_{\frac{1}{2}}$ electron,

$$A_1=\tfrac{2}{3}\mu_0 g(0)M_1; \qquad \text{(A14)}$$

and for a $p_{\frac{1}{2}}$ electron,

$$A_3=-(4/35)\mu_0 g(0)M_3. \qquad \text{(A15)}$$

This last result is identical with the evaluation given by Casimir and Karreman[23] in their original investigation of the octupole interaction in iodine.

For the calculation of second order effects between doublet states, the forms (A7), (A8) of the dipole and quadrupole operators are used. Assuming that both doublet states have identical radial wave functions, the final result is just Eq. (42) with $\xi=\eta=1$.

[23] H. B. G. Casimir and G. Karreman, Physica 9, 494 (1942).

PHYSICAL REVIEW VOLUME 97, NUMBER 2 JANUARY 15, 1955

Shape of Collision-Broadened Spectral Lines*

E. P. Gross†
Laboratory for Insulation Research, Massachusetts, Institute of Technology, Cambridge, Massachusetts

(Received August 24, 1954)

Van Vleck and Weisskopf and Fröhlich have derived a microwave line shape by studying the interruption by collisions of the motion of a classical oscillator. They assume that after the instantaneous impact the oscillator variables are distributed according to a Boltzmann distribution appropriate to the value of the applied field at collision. In contrast to the earlier theory of Lorentz, they obtain the correct static polarization. The procedure involves an assumption of very large velocity during collision. This is criticized on the grounds that the duration of collision is short compared to the resonant period and energy exchanges are of the order of kT. We have derived a line-shape formula assuming that the positions are unchanged after impact. Two extreme models are studied. In one, the oscillators have a Maxwellian distribution of velocities after impact; the second is a Brownian motion treatment. The resulting line shape in both cases is that of a friction-damped oscillator. For collision frequency much less than the resonant frequency, the polarization postulated by the above authors is reached as a result of kinematic motion between collisions, and the line shapes agree. However, to obtain equal line widths and peak absorptions, the collision frequency is twice as large for the present theory. For collision frequency comparable to resonant frequency a less distorted line shape results. For testing the theories, experiments on foreign-gas broadening in the microwave region at pressures of the order of an atmosphere are required. Differences between the theories are small for conditions accessible experimentally at present.

1. INTRODUCTION

THE theoretical determination of the shape of a spectral line, broadened by interactions between the radiating molecule and other systems, is an exceedingly complicated problem. The general case involves a study of the types of interaction possible, treatment of the exchange of energy between internal degrees of freedom and translational motions, questions of coherence, of radiation, etc. In addition, for broad lines one may encounter the characteristic complexities of many-body problems. A clear understanding of the physical processes involved has been gained only in certain limiting cases. There, the consideration of simple models has been useful in calling attention to the ingredients which must enter into more general treatments. The present paper deals with some models which shed light on the processes responsible for the shapes of the spectral lines in gases (chiefly rotational), in the microwave region.

For microwave wavelengths, the energy $\hbar\omega_0$, corresponding to a spectral line of angular frequency ω_0, is usually small compared to the thermal energy kT. This implies that collision-induced transitions between states are important. Indeed, saturation measurements indicate that most collisions involve energy exchanges between the rotational and translational degrees of freedom. If consideration is restricted to foreign-gas broadening (thus excluding the long-range resonance forces), the duration of collision is short compared to the resonant period of the line. It is then useful to introduce for each line a quantity τ, which measures the time between those collisions involving exchanges of energy between translational motions and the relevant internal states. In treatments less schematic than the ones with which we deal, τ is computed in terms of the intermolecular forces. This question is not discussed here; the present work deals with the analysis of some kinetic-statistical aspects of the line-broadening problem. It is of course somewhat arbitrary to split up the problem in

* Sponsored by the U. S. Office of Naval Research, the Army Signal Corps, and the Air Force.

† Present address: Department of Physics, Syracuse University, Syracuse, New York, where the writing of this paper was completed under an Air Force contract.

this manner; a complete treatment would, however, be involved.

Our considerations are in the spirit of previous work by Van Vleck and Weisskopf[1] and also Fröhlich.[2] These treatments are most appropriate for foreign-gas broadening when the nonradiating molecules are more numerous than the radiating ones and form an inert reservoir. It is then permissible to neglect effects which may occur in a purely polar gas, and which have to do with definite correlations between the radiations of two colliding molecules. In addition, the problem of representing the irreversible nature of the absorption process is made simple by the presence of the nonradiating reservoir. In practice correlation considerations probably lead to new results only for highly compressed polar gases. The probable effect of a collision on the state of the radiating molecule is, in general, dependent on the mass ratio of radiating and nonradiating molecules, the type of interaction, the nature of the dynamic order of the reservoir molecules, etc. Study of these factors is important for the understanding of the differences in the dielectric behavior of compressed gases and liquids. Some of these questions will be treated in a communication to follow. For the present we limit ourselves to a criticism and analysis of certain general features of previously proposed models of the collision process. We discuss several alternative approaches. One is a strong-collision model which is perhaps appropriate when radiating and nonradiating molecules have comparable masses, and when, as in a gas, the dynamic ordering of the reservoir molecules is unimportant. The second approach is a Brownian motion treatment of a system of oscillators which is perhaps more appropriate for radiating molecules more massive than those forming the reservoir. More general models can easily be constructed.

In the development of the mathematical theory it has been useful to study the effects of collisions on the absorption or emission of radiation by a classical charge vibrating harmonically. With the help of the concept "virtual oscillator" it is possible to use the classical formulas to find the quantum-mechanical result for a general molecule. This procedure will be followed in the present paper. A more direct quantum treatment, along the lines of the work of Karplus and Schwinger,[3] will be given later. If the absorption and dispersion of electromagnetic radiation is studied in the vicinity of a pure rotation line, four quantities with the dimensions of a frequency enter naturally. These are: the resonant frequency ν_0 of the line; the frequency ν of the applied field; the collision frequency ν_c; and a quantity which is the reciprocal of the duration of collision. The duration of collision is approximately the spatial extent of the interaction divided by the translational velocity, i.e., for foreign-gas broadening, of the order of 10^{-12} sec. The collision duration is thus short compared to the other three periods for microwave rotational lines in the centimeter range.

Van Vleck and Weisskopf[1] have discussed harmonic-oscillator models where collisions are instantaneous, and have derived a line-shape formula [Eq. (22)] which has found wide use in the interpretation of microwave spectra. The line-shape formula has, however, not been adequately checked by experiment when the line width is comparable to the resonant frequency of the line, i.e., at pressures of the order of an atmosphere. Extensive published results exist only for pure NH_3 and ND_3. The ammonia case is complicated by line shifts arising from statistical broadening and inversion splitting, and does not provide a test of the Van Vleck-Weisskopf line shape under the conditions upon which the derivation is based. It is therefore desirable that experiments be performed on the line broadening of simple molecules (e.g., COS) by foreign gases at pressures in the vicinity of an atmosphere. The present paper deals with harmonic oscillators and also makes the assumption of instantaneous collisions. It is believed that the model of the collision process presented here is more realistic for gas broadening. One is led to an alternative line-shape formula which is simply the spectrum emitted by a radiating oscillator moving under the influence of a viscous resistance. The line shape agrees with that of reference 1 when the collision frequency is small compared to the resonant frequency. The discussion of Sec. 4 indicates that the differences between the formulas in the high-pressure region may be measurable, in spite of the contributions of neighboring lines.

2. ASSUMPTIONS OF EARLIER TREATMENTS

The theory of the absorption of energy by a harmonically bound particle is well known. Let e be the charge and m the mass of the particle, ω_0 the resonant angular frequency, $E_0e^{i\omega t}$ the external field and $-mr\dot{x}$ the friction force. Then the steady-state polarization $P(t)$ is given by Eq. (20c) with the resistance $r=1/\tau$. There are three limiting cases of interest. If ω tends to zero, one finds the correct static polarization for a system of N oscillators, i.e., $P=Ne^2E/m\omega_0^2$. All reference to r has disappeared; the static polarization is independent of the rate and mechanism of approach to equilibrium. A second case is r small compared to ω and ω_0; the well-known anomalous dispersion and absorption line shapes result. For r comparable to ω and ω_0 the absorption line shape is asymmetrical and possesses a high frequency tail; it is however symmetrical when plotted against the logarithm of the frequency. A third limit is obtained by setting the resonant frequency equal to zero. One then obtains the simple theory of conduction and light absorption of electrons in metals or in the ionosphere. To adapt this model to describe the rotational lines of a molecule one assigns virtual oscillator

[1] J. H. Van Vleck and V. Weisskopf, Revs. Modern Phys. **17**, 227 (1945).

[2] H. Fröhlich, Nature **157**, 478 (1946); *Theory of Dielectrics* (Oxford University Press, London, 1949). A theory of substantially the same form was presented earlier by R. Kronig, Physica **5**, 65 (1938).

[3] R. Karplus and J. Schwinger, Phys. Rev. **73**, 1020 (1948).

quantities e_{ij}, ω_{ij}, r_{ij} to each transition from level i to level j. At low foreign-gas pressures, i.e., $r_{ij} \ll \omega_{ij}$ and ω, the rotation lines are sharp, well separated, and practically symmetrical. When r_{ij} is comparable to ω and ω_{ij}, the line shape is asymmetric. For larger values one encounters overlap of lines; finally, the instantaneous collision assumption becomes invalid.

Our aim is to derive formulas for the polarization on the basis of more realistic collision models and to discuss the above cases. The principal result is that the simplest collision models give the same results as the friction-damped oscillator. For many molecules there is a range of foreign-gas pressures where individual rotation lines are still separable but where the theories predict different line shapes. In addition, different predictions in the region of strong overlap can in principle be compared with experimental absorption measurements as a function of frequency, pressure, and temperature.

Lorentz attempted to describe the process of interruption by impacts in a more physical manner. While his developments dealt with the optical region, the assumption as to the effects of an instantaneous collision can be adopted to the energy-transfer type of theory of reference 1. The key assumption is that the oscillators which suffer collisions at the same time are redistributed with positions and velocities weighted according to a Boltzmann distribution appropriate to the field-free Hamiltonian $H_0 = p^2/2m + m\omega_0^2 x^2/2$, where p and x are the momentum and position of an oscillator. Lorentz's expression for the polarization is Eq. (20a), which differs from Eq. (20c) by the presence of the term $1/\tau^2$ in the denominator. For collision frequencies much less than ω and ω_0, the difference is negligible. However, as $\omega \to 0$ one obtains the incorrect polarization $P = Ne^2E/m(\omega_0^2 + 1/\tau^2)$ and, as $\omega_0 \to 0$, one fails to obtain free-particle behavior.

The important work of Van Vleck and Weisskopf,[1] and (independently) of Fröhlich,[2] reopened the question of the proper treatment of collisions. These authors attempted to construct a theory which correctly describes line shapes in the microwave region where ω, ω_0, and $1/\tau$ are comparable. To overcome the inadequacies of the Lorentz theory, they assumed that after collision at time t, the oscillator variables are distributed according to a Boltzmann distribution appropriate to the instantaneous Hamiltonian $H(t) = p^2/2m + m\omega_0^2 x^2/2 - eE_0 x \cos\omega t$. This assumption leads to the polarization of Eq. (20b), which is correct in the static limit. However, as $\omega_0 \to 0$ one does not obtain the polarization of a free electron gas, as is the case with Eq. (20c). In fact, the polarization tends to an infinite result. Van Vleck and Weisskopf justified their line shape by finding the quantum-mechanical adaptation of Eq. (20b) for the case of permanent dipoles. In the limit of an infinite moment of inertia, i.e., zero resonant frequency, they obtained the Debye formula,

$$P = \frac{N\mu^2}{3kT} \Big/ (1 + i\omega\tau),$$

where μ is the permanent dipole moment of the molecule. This was considered to be a necessary result of any correct theory. The quantum-mechanical adaptation of formula (20c) does not lead to this result. Formula (20b) has the additional consequence that the absorption tends to the finite value $\omega_p^2/\omega_0^2 \tau c$ as $\omega \to \infty$; the integrated absorption over all frequencies is therefore infinite.

The experimental evidence for the Debye formula lies in dielectric relaxation experiments on liquids. Extensive measurements of the relaxation of dipoles in compressed gases have not been published.[4] The two cases differ in the nature of the dynamic order, and it is probably to the compressed gas that the theories under consideration most nearly refer. That the liquid and compressed gas are profoundly different in their dynamical behavior can be seen, for example, from the fact that the parameter τ (if related to static viscosity) has entirely different dependences on temperature and pressure. Furthermore, measurements[5] of supersonic absorption of compressed gases and liquids near the critical temperature indicate distinct mechanisms for the two cases. Thus we do not consider the Debye law to be a necessary result. This question will be discussed further from the theoretical point of view in a later communication. It will be shown that the Debye relaxation time is essentially different from the time between collisions, τ.

A significant advance in the understanding of the physical assumptions involved in the approaches leading to Eq. (20b) was taken by Van Vleck and Margenau.[6] They pointed out that work is done by the electromagnetic field on the oscillator during an adiabatic collision ("impulsive work"). This is indeed necessary in a theory in which collisions induce amplitude changes since the peak amplitude after collision can differ from that before only if work has been done during the collision. The total energy absorbed from the field is the sum of the work done between and during collisions. However, a less satisfactory implication of the theory is that the oscillator velocity is instantaneously infinite for an infinitesimally short collision. This is necessary so that the mean oscillator position jumps a finite amount to reach the value appropriate to the assumed Boltzmann distribution. The present calculation is based on the conviction that this feature of the theories is not reasonable. The following arguments lead to an alternative model.

If harmonic oscillators collide with the molecules of a buffer medium, one may assume that the kinetic energies after collision have a Maxwellian distribution, since the buffer molecules have such a distribution. This procedure is followed in all the theories dealt with here

[4] Experiments by C. S. E. Phillips of the Laboratory for Insulation Research, Massachusetts Institute of Technology, are in progress.

[5] H. D. Parbrook and E. G. Richardson, Proc. Phys. Soc. (London) **B65**, 437 (1952).

[6] J. H. Van Vleck and H. Margenau, Phys. Rev. **76**, 1211 (1949).

(with the exception of the Brownian motion treatment of Sec. 6). One tacitly assumes that the oscillator mass is of the same order of magnitude as the mass of the buffer molecules and that collisions are strong. Consider, on the other hand, the distribution of positions after collision. We shall argue that a reasonable assumption is that the oscillator does not alter its position during a collision.

A typical energy exchange is of the order of kT. If the duration of the collision is written as β/ω_0, where β is a very small fraction, the distance the oscillator moves during a collision is $\simeq(kT/m)^{\frac{1}{2}}\beta/\omega_0$. In thermal equilibrium the mean square displacement of the oscillator is given by $m\omega_0^2\bar{x}^2/2\simeq kT$ or $(\bar{x}^2)^{\frac{1}{2}}=(2kT/m)^{\frac{1}{2}}/\omega_0$. At collision the oscillator thus moves only a small fraction of a typical amplitude. Furthermore, this distance is independent of the field E_0, so that there can be no tendency to take up a distribution of positions determined by the external field. The action of the applied field during a collision can move the oscillator a distance of the order of $\frac{1}{2}(eE_0/m)(\beta/\omega_0)^2$. This is β^2 times the mean displacement in a static field E_0. It therefore appears that a more reasonable assumption is that the oscillator position is unchanged as a result of collision. The impacts, however, change the velocity instantaneously, implying infinite acceleration. There is impulsive momentum change (force×time) and impulsive work (force×time×particle velocity). Changes in position occur only between collisions as a result of the motion of the particle under the influence of its binding force and the external field. This picture is closely related to the description of collisions in the kinetic theory of gases. In the next section, we develop a mathematical theory embodying this idea. One finds the formula (20c) for the polarization. From the present point of view the reason for the adequacy of the Van Vleck-Weisskopf and Fröhlich formulas when $1/\tau\ll\omega_0$ and ω is the following. Consider a set of oscillators which have made collisions at the same time. The motions will then be rigorously determined by the initial conditions after collision and by the equation of motion containing the binding force and the external field. The electric field distorts the oscillator motion. If $1/\tau\ll\omega_0$ and ω, and furthermore $\omega_0>\omega$, macroscopic quantities such as the polarization have time to approach the value postulated by Van Vleck-Weisskopf and Fröhlich between collisions.[7] No serious error is then made by assuming that the distribution in position is the Boltzmann distribution appropriate to the instantaneous value of the applied field. When $1/\tau$ is comparable to ω_0 or ω, and $\omega\gg\omega_0$, the assumption is inadequate.

3. MATHEMATICAL FORMULATION FOR HARMONIC-OSCILLATOR BROADENING

The one-dimensional harmonic oscillator is characterized by a charge e, a mass m, and a resonant frequency ν_0. The state of the oscillator at time t is prescribed by its position x and its velocity v. We deal with a distribution function $f(x,v,t)$,[2] which determines the statistical behavior of the system of oscillators. The physical interpretation is that $f(x,v,t)dxdv$ gives the number of oscillators at time t with velocities in the range v, $v+dv$ and with positions between x and $x+dx$. If the distribution function is known, it is possible to compute various other quantities of physical interest. The total number of molecules per unit volume (of ordinary, *not* x space) is N, and is independent of time. The distribution function is therefore always required to satisfy

$$N=\int_{-\infty}^{+\infty}\int f(x,v,t)dxdv. \tag{1}$$

The number of oscillators per unit volume at x and time t, is

$$n(x,t)=\int_{-\infty}^{+\infty} f(x,v,t)dv. \tag{2}$$

The mean electric polarization of the oscillators is

$$P(t)=e\int_{-\infty}^{+\infty}\int xf(x,v,t)dxdv. \tag{3}$$

The mean momentum $g(x,t)$ and kinetic energy $k(x,t)$ at x and t are given, respectively, by

$$g(x,t)=m\int_{-\infty}^{+\infty} vf(x,v,t)dv,$$

$$k(x,t)=\frac{m}{2}\int_{-\infty}^{+\infty} v^2f(x,v,t)dv.$$

We now construct the equation governing the behavior of the distribution function. $f(x,v,t)$ changes as a result of purely kinematic action and because of collisions. We write

$$\frac{\partial f}{\partial t}+v\frac{\partial f}{\partial x}+a\frac{\partial f}{\partial v}=\frac{\delta f}{\delta t}\bigg|_c. \tag{4}$$

The term $v\partial f/\partial x$ represents the kinematic drift in x space, and the term $a\partial f/\partial v$, the drift in velocity space as a result of the acceleration a. The expression $\delta f/\delta t|_c$ represents the effects of collisions on the distribution function and will be discussed in detail later. To compute the acceleration we note that the Hamiltonian of the oscillator in an external electric field $E_0\cos\omega t$ is:

$$H(t)=p^2/2m+m\omega_0^2x^2/2-exE_0\cos\omega t. \tag{5}$$

We find the acceleration

$$a=dp/mdt=-\partial H/m\partial x=-\omega_0^2x+eE_0\cos\omega t/m.$$

[7] See reference 2, p. 63. The subsequent development, however, does not make correct use of this insight.

The kinetic equation for the distribution function is therefore

$$\frac{\partial f}{\partial t}+v\frac{\partial f}{\partial x}-\omega_0^2 x\frac{\partial f}{\partial v}+\frac{eE_0\cos\omega t}{m}\frac{\partial f}{\partial v}=\left.\frac{\delta f}{\delta t}\right|_c. \quad (6)$$

In the absence of an electric field the equilibrium distribution is

$$f_0=C_0\exp[-(mv^2+m\omega_0^2x^2)/2kT]; \quad (7)$$

T is the temperature of the reservoir and k is Boltzmann's constant. Since we require $\int\int f_0 dxdv=N$, we find

$$C_0=N\omega_0 m/2\pi kT. \quad (8)$$

The equilibrium distribution in position is

$$n_0(x)=\int f_0 dv=N\omega_0(m/2\pi kT)^{\frac{1}{2}}\exp\left(-\frac{m\omega_0^2x^2}{2kT}\right). \quad (9)$$

If the electric field at time t, $E_0\cos\omega t$, were frozen, the equilibrium distribution which would be reached eventually is

$$f_{eq}=C\exp[-(mv^2+m\omega_0^2x^2-2eE_0x\cos\omega t)/2kT]. \quad (10)$$

One again requires $\int\int f_{eq}dxdv=N$. We shall be concerned with effects which are linear in the electric field strength. To this approximation $C=C_0$. More generally

$$\frac{N}{C}=\frac{2\pi kT}{m\omega_0}\exp\left(\frac{e^2E_0^2\cos^2\omega t}{2kTm\omega_0^2}\right). \quad (11)$$

Let us now consider the collision term $\delta f/\delta t|_c$. Three expressions will be studied here. They are:

(1) The Lorentz assumption,

$$\left.\frac{\delta f}{\delta t}\right|_c=-\frac{1}{\tau}(f-f_0); \quad (12a)$$

(2) the instantaneous equilibrium assumption of Van Vleck and Weisskopf and of Fröhlich,[8]

$$\left.\frac{\delta f}{\delta t}\right|_c=-\frac{1}{\tau}(f-feq); \quad (12b)$$

(3) the assumption of the present report,

$$\left.\frac{\delta f}{\delta t}\right|_c=-\frac{1}{\tau}\left[f-n(x,t)\left(\frac{m}{2\pi kT}\right)^{\frac{1}{2}}\exp\left(-\frac{mv^2}{2kT}\right)\right]; \quad (12c)$$

τ is a constant independent of velocity, representing the time between collisions. We find from the kinetic equation for the above three cases that if f is normalized at any given time

$$\frac{\partial N}{\partial t}=\int\int\left.\frac{\delta f}{\delta t}\right|_c dxdv=0.$$

Thus the normalization condition is maintained by the motions.

The physical content of the Lorentz assumption is that collisions make the distribution relax in a time τ to the equilibrium in the absence of a field. The Lorentz collision term is inadmissible since it leads to an incorrect equilibrium distribution for the case of a static electric field. Collisions fail to maintain the correct distribution. Assumption (2) yields the correct value of the distribution function for static fields. Left and right hand sides of the kinetic equation are separately equal to zero when $f=feq$ for static fields. The same statement holds when assumption (3) is employed. The physical content of assumption (2) is that collisions tend to make the distribution relax in a time τ to the equilibrium value characteristic of the instantaneous field at time t. This content can be expressed in another way. The number of molecules ejected from the range $x, x+dx$ and $v, v+dv$ is proportional to the number actually present in this range at time t; this is the $-f(x,v,t)/\tau$ term. The number per second re-emitted into the range is given by feq/τ. The mathematical form of assumption (3) is somewhat different from the others; it contains the integral quantity $n(x,t)=\int f dv$. Assumption (3) states that the number per second reemitted into the velocity range $v, v+dv$ as a result of collisions has the Maxwellian form. The number re-emitted into the spatial region, $x, x+dx$, is proportional to the total number in that region before collisions.

The difference between assumption (3) and assumptions (1) and (2) may be seen by computing $\int\delta f/\delta t|_c dv$. This expression represents the change per unit time in the number of particles at x, t. It vanishes under assumption (3) but not under assumptions (1) and (2). The change per unit time due to collisions of both kinetic energy and total energy is nonzero for all three assumptions. This is indeed necessary to represent the degradative nature of the absorption process.

4. THEORY OF ABSORPTION AND DISPERSION

We investigate only processes linear in the electric field strength. Let us adopt complex notation and write

$$E=E_0\exp(i\omega t).$$

We introduce the dimensionless quantities $\phi(x,v,t)$ and $\nu(x,t)$ defined by

$$f=f_0(1+\phi), \quad (13)$$

$$n=n_0(1+\nu); \quad (14)$$

ϕ and ν are proportional to the electric field strength and are small compared to unity. If terms containing products of ϕ, ν, or E are neglected, one obtains the

[8] Fröhlich uses the language of distribution functions, Van Vleck and Weisskopf that of mean-free-path theory. The latter follows the history of an individual particle and is intimately connected with the characteristics of the partial differential equation satisfied by the distribution function.

kinetic equation,

$$\frac{\partial\phi}{\partial t}+\frac{1}{\tau}\phi+v\frac{\partial\phi}{\partial x}-\omega_0^2x\frac{\partial\phi}{\partial v}=\frac{eE_0\cos\omega t v}{kT}+\Delta. \quad (15)$$

Δ takes on the following values for the three assumptions discussed above

$$(1)\quad \Delta=0, \quad (16a)$$

$$(2)\quad \Delta=(1/\tau)(eEx/kT), \quad (16b)$$

$$(3)\quad \Delta=\nu/\tau. \quad (16c)$$

In the present section steady-state solutions will be studied. ϕ and ν are assumed to have the form $\phi=\phi_0\exp(i\omega t)$, $\nu=\nu_0\exp(i\omega t)$. For the discussion of this section we seek solutions of the form[2]

$$\phi=-(\alpha x+\beta v)E_0\exp(i\omega t). \quad (17)$$

The values of α and β are found on inserting the expression for ϕ in the kinetic equation and equating the coefficients of x and v separately to zero. With the chosen form of ϕ the value of ν is

$$\nu=\left(\int\phi f_0dv\Big/\int f_0dv\right)=-\alpha xE_0\exp(i\omega t). \quad (18)$$

The three values of ϕ are

$$(1)\quad \phi=\frac{eE_0}{kT}e^{i\omega t}\frac{1}{(1/\tau+i\omega)^2+\omega_0^2}\times\left[\omega_0^2x+\left(\frac{1}{\tau}+i\omega\right)v\right], \quad (19a)$$

$$(2)\quad \phi=\frac{eE_0}{kT}e^{i\omega t}\frac{1}{(1/\tau+i\omega)^2+\omega_0^2}\times\left[\left(\omega_0^2+\frac{1}{\tau^2}+\frac{i\omega}{\tau}\right)x+i\omega v\right], \quad (19b)$$

$$(3)\quad \phi=\frac{eE_0}{kT}e^{i\omega t}\frac{1}{\omega_0^2-\omega^2+i\omega/\tau}[\omega_0^2x+i\omega v]. \quad (19c)$$

The polarization is $P=e\int\int xf_0\phi dxdv$ and depends only on the coefficient α. One finds

$$(1)\quad P=\frac{Ne^2E_0}{m}\frac{e^{i\omega t}}{(1/\tau+i\omega)^2+\omega_0^2}, \quad (20a)$$

$$(2)\quad P=\frac{Ne^2E_0}{m\omega_0^2}\frac{[\omega_0^2+1/\tau^2+i\omega/\tau]e^{i\omega t}}{(1/\tau+i\omega)^2+\omega_0^2}, \quad (20b)$$

$$(3)\quad P=\frac{Ne^2E_0}{m}\frac{e^{i\omega t}}{\omega_0^2-\omega^2+i\omega/\tau}. \quad (20c)$$

The complex dielectric constant ϵ may be written as $\epsilon=\epsilon'-i\epsilon''$. It is obtained from the formula $\epsilon-1=4\pi P/E$. The absorption, a, is given by $a=\omega\epsilon''/c$, where c is the velocity of light.

The expressions for ϵ', ϵ'', a are (with $\omega_p^2=4\pi Ne^2/m$),

$$(1)\quad \epsilon'-1=\omega_p^2(1/\tau^2+\omega_0^2-\omega^2)/D, \quad (21a)$$

$$\epsilon''=\omega_p^2 2\omega/\tau D, \quad (21b)$$

$$a=2\omega^2\omega_p^2/c\tau D; \quad (21c)$$

$$(2)\quad \epsilon'-1=\omega_p^2\left[\left(\omega_0^2+\frac{1}{\tau^2}\right)\times\left(\omega_0^2+\frac{1}{\tau^2}-\omega^2\right)+\frac{2\omega^2}{\tau^2}\right]\Big/\omega_0^2D, \quad (22a)$$

$$\epsilon''=\omega_p^2\omega\left(\omega_0^2+\frac{1}{\tau^2}+\omega^2\right)\Big/\omega_0^2\tau D, \quad (22b)$$

$$a=\omega_p^2\omega^2\left(\omega_0^2+\frac{1}{\tau^2}+\omega^2\right)\Big/\omega_0^2\tau cD,[10] \quad (22c)$$

where the denominator D is: $D=(1/\tau^2+\omega_0^2-\omega^2)^2+4\omega^2/\tau^2$;

$$(3)\quad \epsilon'-1=\omega_p^2(\omega_0^2-\omega^2)/(\omega_0^2-\omega^2)^2+\omega^2/\tau^2, \quad (23a)$$

$$\epsilon''=\frac{\omega_p^2\omega}{\tau}\Big/(\omega_0^2-\omega^2)^2+\omega^2/\tau^2, \quad (23b)$$

$$a=\frac{\omega_p^2\omega}{\tau c}\Big/(\omega_0^2-\omega^2)^2+\omega^2/\tau^2. \quad (23c)$$

Let us now compare the predictions of the theories (2) and (3). There is a variety of cases, the exact situation depending on the relations between ω, ω_0, and $1/\tau$.

Consider, as a first case, the low-frequency wing of the line. For ω much less than both $1/\tau$ and ω_0, the absorption according to (2), is

$$a=\omega_p^2\omega^2/\omega_0^2\tau c\left(\omega_0^2+\frac{1}{\tau^2}\right).$$

According to (3), $a=\omega_p^2\omega^2/c\tau\omega_0^4$. The results agree if $1/\tau\ll\omega_0$. This has an important consequence if one is considering the absorption in the microwave region at high pressures, when there are appreciable contributions from overlapping lines of high resonant frequency. For most of these lines, $1/\tau\ll\omega_0$, so that the contributions to the absorption are the same for theories (2) and (3). Of course, if the collision frequency is so great that it is comparable to the frequency of the line of maximum intensity [for a rotator of moment of inertia I, a fre-

[10] This line shape was derived by R. Kronig, Physica **5**, 65 (1938).

quency of $(kT/I)^{\frac{1}{2}}$] there will be important differences. Theory (2) then predicts a considerably smaller absorption than does (3). In addition, at a fixed frequency ω, the increase in absorption as $1/\tau$ increases is less rapid according to (2) than according to (3). These effects are, however, somewhat obscured by the increasing untrustworthiness of the instantaneous collision hypothesis.

As a second case, we examine the behavior in the high-frequency wing of the line. Theory (2) yields, when $\omega \gg 1/\tau$, ω_0, the finite absorption $a=\omega_p^2/\omega_0^2\tau c$. Theory (3) yields zero for the absorption. Thus, unless a cutoff is introduced because of the failure of the instantaneous collision assumption, theory (2) leads to an infinite absorption when integrated over all frequencies. In practice, observation of the high frequency tail of one of the low lying rotation lines is difficult. Overlapping of higher rotational lines is important at collision frequencies of the order of ω_0. It is only for such collision frequencies that the predictions of theories (2) and (3) differ significantly. Since the collision frequency may not be strictly proportional to foreign-gas pressure, and collision frequencies for higher lines differ in general from those for the lower lines, it is probably no easy task to find clearcut verifications of any theory.

A third case is $1/\tau \ll \omega$ and ω_0. This will occur at low pressures when one finds narrow lines. Confining oneself to the body of the line, where $\omega_0^2-\omega^2$ is of the same order of magnitude as $1/\tau^2$, we find $D \simeq (\omega_0^2-\omega^2)^2+4\omega^2/\tau^2$. The absorption according to theory (2) is

$$a=\frac{\omega_p^2 2\omega^2}{\omega_0^2\tau c}\Big/\left[(\omega_0^2-\omega^2)^2+4\omega^2/\tau^2\right].$$

This has the same shape as Eq. (23c); however, to obtain the same line width the collision frequency $1/\tau$ of assumption (3) must be taken as twice that entering in assumption (2). This is an important point, since a more detailed theory expresses the collision time τ in terms of molecular properties such as dipole moment, quadrupole moment, polarizibility. To compare with experiment one must obtain τ from observed line widths. It is then relevant whether the observed line width should be identified with $1/\tau$ or $2/\tau$.

For the third case (narrow lines) the peak of the absorption occurs at approximately $\omega=\omega_0$ on both theories (2) and (3). Theory (2) predicts the maximum absorption $a(\omega=\omega_0)=\omega_p^2\tau/2c$ while theory (3) yields twice this value. If the collision frequency of (3) is chosen to make the line width agree with that of (2) the absolute intensities will also be the same. However, comparing with the first case we see that the absorption in the low-frequency wing will be twice as large for theory (3).

Fourth, we consider the situation where ω, ω_0, and $1/\tau$ are all of comparable magnitude. Formula (23c) has the important properties that the peak of the absorption always occurs at $\omega=\omega_0$ and that the absorption is symmetric if one plots against the natural logarithm of the frequency. According to formula (22c) the position of the maximum absorption shifts to higher frequencies as $1/\tau$ becomes comparable to ω_0. The position of maximum absorption is given by

$$\omega_{\max}^2=\frac{(\omega_0^2+1/\tau^2)^2+\left[(\omega_0^2+1/\tau^2)^4+(\omega_0^2+1/\tau^2)^3(3\omega_0^2-1/\tau^2)\right]^{\frac{1}{2}}}{3\omega_0^2-1/\tau^2}.$$

For $1/\tau=0.3\omega_0$, this shift is approximately 10 percent; for $1/\tau=0.5\omega_0$ it is 25 percent. The behavior of the maximum of the line is perhaps the most promising way of testing the theories. We must however notice that at $1/\tau=0.3\omega_0$ the contributions of overlapping lines are also of the order of 10 percent. To add to the difficulties the contributions from the low-frequency wings of the higher rotational lines for the values of $1/\tau$ chosen, are likely to be larger according to formula (23c). This also leads to a shift towards the high frequency side. For comparison of theory and experiment the situation is more favorable for the third or fourth rotational line rather than the first. The ratio of statistical weights of the higher lines to the given line is then less, making the effects of higher lines of less importance. At $\omega=\omega_0$ the absorption according to Eq. (23c) is $a=\omega_p^2\tau/c$ aside from the effects of overlapping lines, it decreases with foreign-gas pressure. Formula (22c) has an additional factor which partially counteracts the decrease.

Finally, we examine the limit $\omega_0 \ll \omega$ and $1/\tau$, i.e., resonant frequency small compared to applied frequency and collision frequency. Such a case is difficult to realize for the rotation spectrum without the crucial intervention of overlapping lines. For $\omega_0=0$, assumption (3) yields the complex dielectric constant $\epsilon-1=\omega_p^2/(i\omega/\tau)-\omega^2$, which is characteristic of a free-electron gas. Both the real part of the dielectric constant, and the absorption $a=\omega_p^2/c\tau(\omega^2+1/\tau^2)$ are finite. The assumption (2) of Van Vleck and Weisskopf, on the other hand, leads to

$$\epsilon-1\rightarrow\frac{\omega_p^2}{\omega_0^2\tau}\Big/\left(\frac{1}{\tau}+i\omega\right),\quad a=\frac{\omega_p^2\omega^2\tau}{\omega_0^2 c}\Big/(1+\omega^2\tau^2).$$

Both expressions become infinite as $\omega_0 \rightarrow 0$.

5. INITIAL VALUE PROBLEM

In the preceding section, we have investigated solutions which are of the type $\phi=-(\alpha x+\beta v)E(t)$. In addition, the study has been restricted to the response to a monochromatic applied field. Let us now briefly examine, within the framework of solutions of this type, initial value problems, e.g., the decay of the polarization or distribution function to equilibrium. Writing

$$\phi=A(t)x+B(t)v, \tag{24}$$

one finds, on setting coefficients of x and v separately to

zero,

$$\frac{dA}{dt}+\frac{1}{\tau}A-\omega_0^2B=\Delta/x, \tag{25a}$$

$$\frac{dB}{dt}+\frac{1}{\tau}B+A=\frac{eE(t)}{kT}. \tag{25b}$$

Consider a problem in which the distribution is Maxwellian at $t<0$, i.e., $\phi=0$, so that $A=B=0$. At $t=0$, a static field $E(t)$ is applied. Equation (15) has unique solutions if the distribution function $\phi(x,v,0)$ is prescribed at $t=0$. Thus solutions of the form $\phi=A(t)x+B(t)v$ are appropriate for those problems in which the initial distribution is also of this form. This is certainly the case when $\phi(x,v,0)\equiv 0$ at $t=0$. The solutions of the problem are

(1)
$$A(t)=\frac{eE_0\omega_0^2}{kT(\omega_0^2+1/\tau^2)}\left[1+\frac{e^{-t/\tau}}{2\omega_0}\left\{\left(\frac{i}{\tau}-\omega_0\right)\times e^{\omega_0 t}-\left(\frac{i}{\tau}+\omega_0\right)e^{-i\omega_0 t}\right\}\right], \tag{26a}$$

$$B(t)=\frac{eE_0}{\tau kT(\omega_0^2+1/\tau^2)}\left[1+\left\{-\frac{(1+i\omega_0 t)}{2}\times e^{i\omega_0 t}-\frac{(1-i\omega_0 t)}{2}e^{-i\omega_0 t}\right\}e^{-t/\tau}\right]; \tag{26b}$$

(2)
$$A(t)=\frac{eE_0}{kT}(1-\cos\omega_0 t)e^{-t/\tau}, \tag{27a}$$

$$B(t)=\frac{eE_0}{kT}\frac{\sin\omega_0 t}{\omega_0}e^{-t/\tau}; \tag{27b}$$

(3)
$$A(t)=\frac{eE_0}{kT}\left(1+\frac{e^{-t/2\tau}}{2(1/\tau^2-4\omega_0^2)^{\frac{1}{2}}}\times\left\{-\frac{1}{\tau}[1+(1-4\omega_0^2\tau^2)^{\frac{1}{2}}]\times\exp\left[+\frac{t}{2\tau}(1-4\omega_0^2\tau^2)^{\frac{1}{2}}\right]+\frac{1}{\tau}[-1+(1-4\omega_0^2\tau^2)^{\frac{1}{2}}]\times\exp\left[-\frac{t}{2\tau}(1-4\omega_0^2\tau^2)^{\frac{1}{2}}\right]\right\}\right); \tag{28a}$$

$$B(t)=\frac{eE_0}{kT}\frac{e^{-t/2\tau}\{\exp[(1/\tau^2-4\omega_0^2)^{\frac{1}{2}}t/2]-\exp[-(1/\tau^2-4\omega_0^2)^{\frac{1}{2}}t/2]\}}{(1/\tau^2-4\omega_0^2)^{\frac{1}{2}}}. \tag{28b}$$

It is seen that with the assumptions (1) and (2) the distribution function and polarization tend to the final values as $e^{-t/\tau}e^{\pm i\omega_0 t}$, i.e., with damped, oscillatory behavior. Assumption (3) leads to a dependence which is a linear superposition of two exponentials. For $1/\tau\ll 2\omega_0$ the decay is exactly as in assumption (2), i.e., as a damped oscillation. [Again the collision frequency must be taken as twice that of assumption (2).] If the collision frequency is so large that $1/\tau>2\omega_0$ the decay is purely exponential. For $1/\tau\gg 2\omega_0$ one of the decay rates is $\omega_0(\omega_0\tau)$, which is much smaller than ω_0. The high collision rate makes it difficult for the oscillator to adjust to the new equilibrium position in space.

Initial value problems of a more general nature than the one treated above may be studied by the method of characteristics. Alternatively, the problem of solving the kinetic Eq. (15) may be attacked by expanding the distribution function in Hermite polynomials. After performing a Laplace transformation with respect to the time variable one is left with simple difference equations. These may be treated by continued fraction methods. We do not undertake this analysis, which yields the exact distribution function, since the coefficient of $\dot{x}$ is unaltered. Thus the simpler approach of Secs. (4) and (5) yields the correct polarization.

6. GENERAL TREATMENTS OF OSCILLATOR BROADENING

The collision model discussed in Secs. (4) and (5) contains two separate assumptions. The first, and most essential assumption was that the position coordinate of an oscillator is unaltered by an instantaneous collision. The second assumption was that the oscillator velocities after impact were redistributed according to a Maxwellian distribution. As mentioned earlier, the second assumption does not hold for a general mass ratio of radiating and nonradiating molecules. If the first assumption is retained, one can write quite generally

$$\left.\frac{\delta f}{\delta t}\right|_c=\int f(v',x,t)A(v',v)dv'-f(v,x,t)\int A(v,v')dv'. \tag{29}$$

$A(v',v)$ is a transition probability which describes the effects of collisions. It is independent of time and the external field, a property which appears in keeping with the inert reservoir character of the nonpolar foreign gas. Our earlier assumption was

$$A(v,v')=A(v')=\frac{1}{\tau}\left(\frac{m}{2\pi kT}\right)^{\frac{1}{2}}\exp\left(-\frac{mv^2}{2kT}\right).$$

Keilson and Storer[11] have considered kernels of the form

$$A(v,v')=\mathfrak{A}(v'-\gamma v), \quad \text{with } 0\leqslant\gamma\leqslant 1. \tag{30}$$

Since

$$\frac{1}{\tau}=\int_{-\infty}^{+\infty}A(v,v')dv'=\int_{-\infty}^{+\infty}\mathfrak{A}(\zeta)d\zeta, \tag{31}$$

[11] J. Keilson and J. E. Storer, Quart. Appl. Math. **10**, 243 (1952).

the collision time is independent of velocity, as has been the case for the theories of earlier sections. The particular kernel studied by Keilson and Storer is

$$A(v,v')=(1/\tau)(\beta/\pi)^{\frac{1}{2}}\exp[-\beta(v'-\gamma v)^2]. \quad (32)$$

Correct equilibrium is obtained with $\beta(1-\gamma^2)=m/2kT$. The case where $\gamma=0$ corresponds to the treatment of Sec. (4). The limit $\gamma\to 1$ is related to the Fokker-Planck equation, for which

$$\left.\frac{\delta f}{\delta t}\right|_c=\beta'\left\{\frac{kT}{m}\frac{\partial^2 f}{\partial v^2}+\frac{\partial}{\partial v}(vf)\right\}. \quad (33)$$

Intermediate values of γ exhibit some features of the scattering of objects of different mass ratios. The complete kinetic equation is obtained by combining Eq. (6) with (33) or Eq. (29) with (32).

As in Secs. (4) and (5), we do not enter into a discussion of the complete solutions. We are interested in the behavior of the polarization, which is e times the first moment in x of the distribution

$$P=e\langle x\rangle=\int_{-\infty}^{+\infty}\int x f dx dv.$$

Equation (33) yields

$$\int\int x\left.\frac{\delta f}{\delta t}\right|_c=0, \quad (34)$$

and

$$\int\int v\left.\frac{\delta f}{\delta t}\right|_c dxdv=-\beta'\langle v\rangle. \quad (35)$$

Forming the x moment of the left-hand side of Eq. (6) and combining with Eq. (34), we have

$$\frac{\partial}{\partial t}\langle x\rangle-\langle v\rangle=0. \quad (36)$$

For the v moment one finds the equation

$$\frac{\partial}{\partial t}\langle v\rangle+\omega_0^2\langle x\rangle-\frac{NeE_0}{m}\cos\omega t=-\beta'\langle v\rangle. \quad (37)$$

The x moment then satisfies the equation

$$\frac{\partial^2\langle x\rangle}{\partial t^2}+\beta'\frac{\partial}{\partial t}\langle x\rangle+\omega_0^2\langle x\rangle=\frac{NeE_0}{m}\cos\omega t. \quad (38)$$

The polarization is the same as Eq. (20c) with $\beta'=1/\tau$. The more general Eqs. (29) and (32) yield

$$\int\int x\left.\frac{\delta f}{\delta t}\right|_c dxdv=0, \quad (39)$$

$$\int\int v\left.\frac{\delta f}{\delta t}\right|_c dxdv=\frac{(\gamma-1)}{\tau}\langle v\rangle. \quad (40)$$

One obtains the friction-damped oscillator behavior again, with a mean time between collisions of $\tau/1-\gamma$. The Fokker-Planck limit is obtained by letting $\gamma\to 1$ and $\tau\to 0$ in such a way that limit $(\gamma-1)/\tau=\beta'$.

The results of the present section indicate that the line shape (20c) is quite general if positions are unchanged by collisions. However, the higher moments of the distribution are sensitive to the particular model chosen. The fact that the first moment (polarization) has this behavior is peculiar to the harmonic oscillator. Other systems, such as rigid dipoles, have polarizations which depend on the precise nature of the velocity exchanges on collision.

ACKNOWLEDGMENT

The author is deeply indebted to Professor A. von Hippel for calling his attention to this problem and for many illuminating discussions. Dr. C. S. E. Phillips and Mr. J. Lebowitz have contributed critical and stimulating comments.

Charge Equilibrium Ratios for Hydrogen Ions from Solids*

JAMES A. PHILLIPS
University of California, Los Alamos Scientific Laboratory, Los Alamos, New Mexico
(Received August 18, 1954; revised manuscript received October 13, 1954)

Experimental values for the ratios of the positive, neutral, and negative components have been obtained for proton beams in the energy region 3–200 kev emerging from Al, Be, Ca, Ag, Au, and SiO foils. Clean surfaces were obtained by evaporating fresh material onto the foil during a run. As expected, the equilibrium charge ratio appears to be determined by the last few (~5) atoms on the exit side of the foil. At 37±3 kev, the positive and neutral components are equal for all materials tested, except Ca (18±2 kev). Significant differences have been found in the charged components from these metals. The observed ratios of the components for different metals parallel their work functions at proton energies >100 kev.

INTRODUCTION

WHEN a beam of low-energy hydrogen ions passes through material, electrons are captured and lost by the ions, and particles of positive, neutral, and negative charge arise. A knowledge of such processes is important in the understanding of the slowing down of the ions by the material. Of interest are the ratios of the component intensities as a function of the ion energy and the dependence on the material through which the beam is passed.

A comprehensive review has been made by Allison and Warshaw[1] of the present knowledge of the charge exchange processes. To avoid repetition the reader is referred to this review for the considerable work done with proton beams passing through gases.

For the case of solids Hall[2] has measured the charged and neutral fractions of an equilibrated beam emergent from four metals in the energy range 20–400 kev. It was found that the equilibrium ratios produced in Be, Al, and Ag foils were the same. A slightly larger neutral component was obtained from the Au foil.

It has also been reported[3] that the charge ratios for helium ions emergent from solids is independent of the stopping material.

It has been suggested[1,2,4] that these results may be due to a surface layer of foreign material. The electron capture and loss cross sections are so large (~10^{-16} cm^2) that contaminating layers only a few atoms thick will completely alter the characteristic equilibrium ratio for the solid.

The present work was undertaken in an attempt to measure the equilibrium charge ratios for a surface as clean as practicable in the target area of an accelerator. Fresh material could be evaporated onto the exit side of the foil during an experiment and the rate of growth of the oxide and foreign material determined by changes in the charge ratios. Al, Be, Ca, Ag, Au, and SiO were evaporated in this manner. The negative component, which until that time was unknown in this energy range, was also measured.

APPARATUS

The apparatus is shown in Fig. 1. Analyzed protons of energy 5–250 kev from a Cockcroft-Walton accelerator passed through a thin aluminum foil. The emergent beam was collimated, magnetically analyzed into its charge components and detected by a phosphor-photomultiplier system. The charged components could be

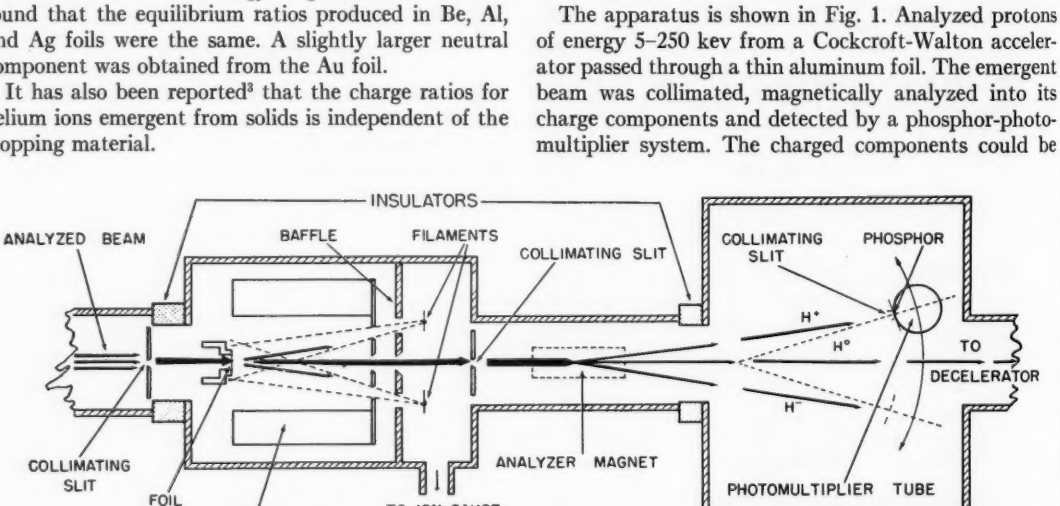


FIG. 1. Schematic diagram of the apparatus, showing the evaporator and detector.

* Work performed under the auspices of the U. S. Atomic Energy Commission.
[1] S. K. Allison and S. D. Warshaw, Revs. Modern Phys. **25**, 779 (1953).
[2] T. A. Hall, Phys. Rev. **79**, 504 (1950).
[3] G. Dissanaike, Phil. Mag. **357**, 1051 (1953).
[4] J. A. Phillips, Phys. Rev. **91**, 455(A) (1953).

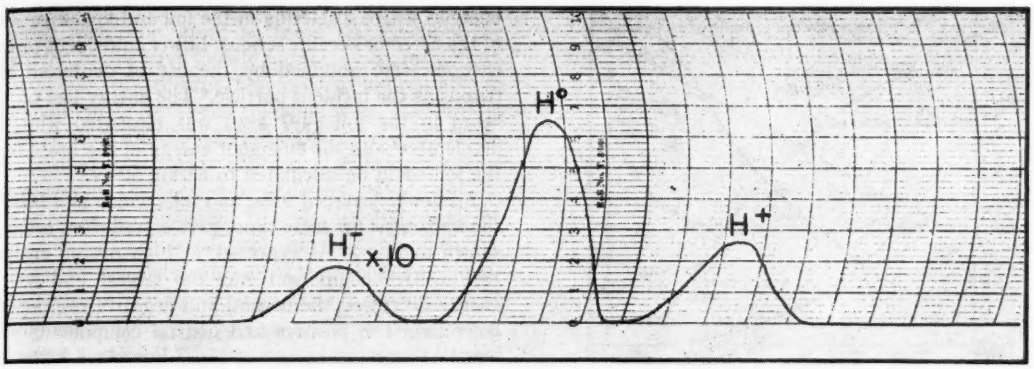


FIG. 2. Record from the phosphor photomultiplier detector, showing the resolved charge components from a foil at an emergent proton energy of 20.5 kev.

sent into a decelerator by which the energy of the emergent beam was measured.

All the foils were of aluminum (~5–10 $\mu g/cm^2$) made by an evaporation technique.[5] The aluminum surface was always on the emergent side of the foil. Condensable matter in the vacuum was excluded from the exit side of the foil by providing liquid air cooled surfaces which surrounded it. Two evaporation filaments were provided by means of which layers of fresh material could be deposited onto the foil. Magnetically operated shutters made it possible to bring the filaments up to temperature before exposing them to the foils.

A 1P21 photomultiplier was mounted at the end of a rotating arm so that the detector could be swept through the three charge components. For complete separation of the three beams, the detector moved through an angle of $\pm 5°$. As single particles were not counted but the integral effect of many ions was recorded, a fast phosphor was unnecessary and the phosphor was chosen for constancy under proton bombardment. Calcium tungstate was used with success and was coated onto a half-aluminized glass plate (to avoid charging up the phosphor) which in turn was mounted on the glass envelope of the photomultiplier. A narrow slit (~0.005 in. wide) was placed immediately in front of the phosphor at right angles to the motion of the detector. In this manner the beam, about 0.20 in. in diameter, would strike the same area of the phosphor as the detector was moved from one angle to another. A Brown recording milliammeter measured the output current of the photomultiplier after passing through an electrometer circuit. Records as shown in Fig. 2 were obtained as the detector automatically swept through the three charged components.

Several tests were necessary to show that the detector was satisfactory.

1. Linearity of detector response to beam current. Beam currents were varied by as much as 10^5; the response of the tube was linear to better than 1 percent over this range.

2. Sensitivity as a function of the angle of the detector. With no foil in place, the analyzer magnet was varied to displace the beam from 0° to ±15°. The response of the detector was constant (~1 percent) up to angles of ±10° and increased thereafter. The components were never separated by more than 5° during the experiment.

3. As the axis of rotation of the detector was not located at the center of the analyzing magnet, it was convenient to show that the intensity of any charge component was proportional to the peak current as measured by the detector; see Fig. 2. Within 0.5 percent, the areas under the current curves were proportional to the peak currents if the detector was not moved through angles greater than ±5°.

4. The total beam was equal to the sum of its components. The response of the phosphor is considered to be identical for the positive, neutral, and negative components of the beam. The charge exchange cross sections are so large that as soon as an ion has traversed the first few atoms of the phosphor it has completely lost its initial identity. The total beam and its three components could be measured separately and within 1 percent the sum was equal to the total unresolved beam.

5. Light produced by the beam passing through the residual gas in the vacuum chamber could be detected by the phosphor. The only response of the detector (when a glass plate was placed immediately in front of the detector) was from the fluorescence of the glass by the beam and was a thousand times smaller than with the same beam current on the phosphor.

6. The distance of the analyzer magnet from the foil was 20 cm and it was important that no appreciable charge exchange should occur in this region. The pressure was increased from 5×10^{-6} mm Hg to 10^{-4} mm Hg and no change in the charge ratio greater than 1 percent was observed.

[5] G. A. Sawyer, Rev. Sci. Instr. 23, 604 (1952).

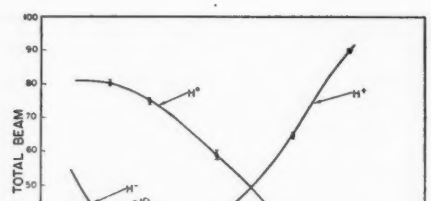


FIG. 3. Equilibrium proton charge ratios from an aluminum foil contaminated with foreign material on its surfaces as a function of exit energy. (Several experimental points are included to show the probable error.)

RESULTS

With an aluminum foil in place, the relative intensities of the three charged components were measured several times at each energy; see Fig. 2. The beam diameter was 3 mm and currents of 0.001–0.3 μa were used, the larger currents being necessary at low beam energies where scattering in the foil and low sensitivity of the detector became serious. Lower equivalent proton energies were obtained by the use of deuterons and tritons as the incident particle.[6] The energy lost by the beam in the foil (4–7 kev) was measured with the decelerator[7] and the emergent energy of the beam from the foil could be calculated to within 50 ev.

All newly installed foils, and foils which had been in the apparatus for some time, gave equilibrium ratios as shown in Fig. 3. As expected, at high proton energies the positive component was the largest and as the energy decreased the neutral and negative components increased. The positive and neutral components were equal at a proton energy of 37±3 kev. At 3.2 kev, the lowest energy examined, the negative component (~5.5 percent) was still smaller than the positive component (~13 percent). The constancy of these ratios for untreated foils was used as a check on the apparatus before each run.

When fresh material was evaporated onto the exit side of the foil the following effects were noted with each metal.

Aluminum

Immediately after the evaporation of a layer of aluminum the ratios were found to have changed as shown in Fig. 4(a). In this case, with a beam energy of 20.5 kev, the H^+ component was found to have changed immediately after the evaporation from 38 percent to 26 percent. The H^+ component was then found to increase and come back to its original value. Decreasing the pressure by a factor of 6 increased this time by about the same factor.

It is thought that initially the surface has a layer of foreign material which gives the 38 percent H^+ component. The new value of the H^+ component (26 percent) immediately after evaporation is due to the fresh aluminum layer on the surface which subsequently either oxidizes in residual air in the vacuum or is covered by a layer of dirt resulting in the initial 38 percent H^+ component. The two possibilities, oxidation or accumulation of dirt, were separated by depositing first a layer of gold followed later by a layer of aluminum under the same conditions. The results are shown in Fig. 4(b). Gold, which does not oxidize appreciably, took several minutes to return to the initial condition, whereas aluminum which does oxidize was relatively much faster. It is concluded that the rapid change of aluminum is due to the process of oxidation and the slow change of gold is due to the growth of a dirt layer.

Some effort was made to measure that thickness of the evaporated aluminum layer in which charge-exchange equilibrium will be reached. It was found that a layer about 50 ev (or 6 atoms) thick is sufficient for equilibrium. It is not known how significant this meas-

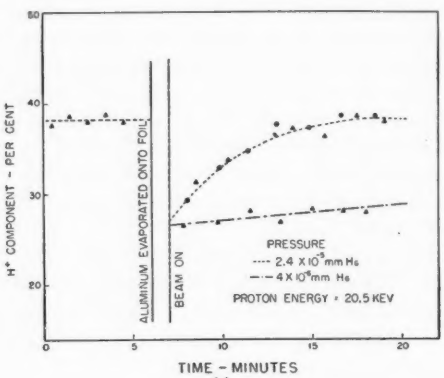


FIG. 4. Changes in the amount of the H^+ component as a function of time as (a) aluminum is evaporated onto the foil, and (b) gold and then aluminum are evaporated under identical conditions.

[6] J. A. Phillips, Phys. Rev. **90**, 533 (1953).
[7] Arnold, Phillips, Sawyer, Stovall, and Tuck, Phys. Rev. **93**, 483 (1954).

urement might be, since at these temperatures the aluminum atoms have mobility and may aggregate into crystals. For this reason it is considered unprofitable to determine more precisely this minimum equilibrium thickness since an area and not a thickness effect might be measured. During the experiment layers of at least five times this thickness were deposited.

It is of course impossible to state definitely that the data, when extrapolated back to the time when the evaporation has been completed, are those for aluminum. Changes (to the aluminum) may have occurred during the evaporation process and the transport of the metal to the foil. Moreover, sixteen seconds elapsed after an evaporation before a measurement of the charge ratios could be made. This possibility is thought, however, to be unlikely, as the extrapolations of data taken with pressures differing by a factor of 6 agree within the experimental error.

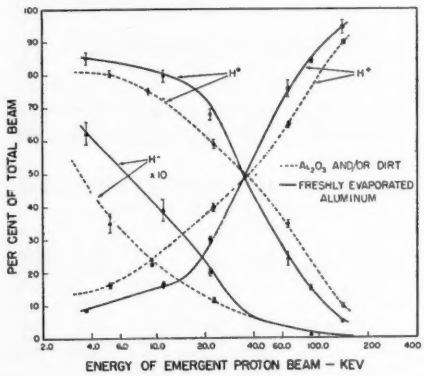


FIG. 5. Comparison of the equilibrium charge ratios from a freshly evaporated aluminum surface with a "dirty" surface.

Evaporating aluminum at several proton energies and extrapolating to zero time, the ratios for aluminum were found as shown in Fig. 5. A few of the experimental points are included to show the probable error.

Beryllium

Beryllium was evaporated from tantalum filaments and found to be very similar in behavior to aluminum. The rate of growth of the oxide made it necessary to evaporate fresh material onto the foil at each energy and to extrapolate the data to zero time.

Silver, Gold, and SiO

The equilibrium ratios could be measured at several energies before the growth of contaminating layers of dirt became serious. The condition of the foil was constantly monitored by repeatedly returning to that energy first used after an evaporation.

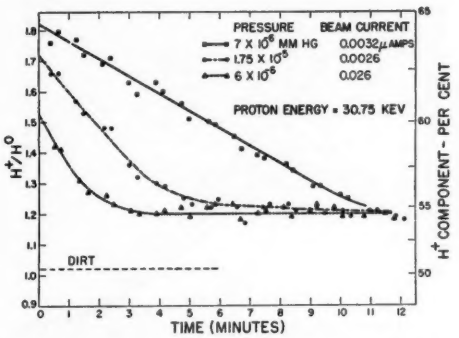


FIG. 6. Changes in the H^+ component from calcium as a function of time under different beam and vacuum conditions.

Calcium

Considerable difficulty was experienced with the high chemical reactivity of this metal. Not only did the metal react quickly on the foil, but the metal reacted while on the evaporating filaments, making it increasingly difficult to evaporate it during a run. Both the vacuum pressure and the magnitude of the beam current through the foil were found to change the rate of growth of the oxide, as shown in Fig. 6. Such data show that at the lowest pressures ($\sim 5 \times 10^{-6}$ mm Hg) and the smallest beam currents (~ 0.001 μa) used in the experiment the surface had probably changed before a measurement of the charge components could be made. The results given for calcium are extrapolations of the data taken under the most favorable conditions but are probably contaminated to some extent.

The results for the metals examined in this experiment are shown in Figs. 7 and 8, and numerical values in Table I.

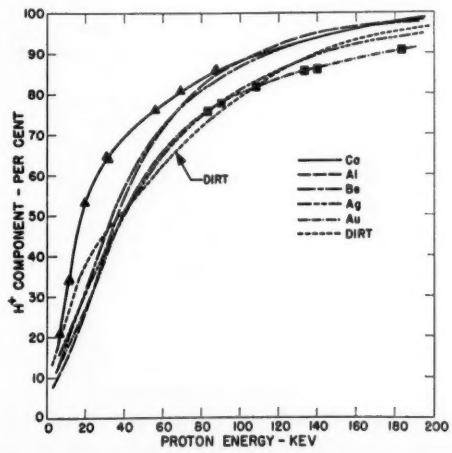


FIG. 7. H^+ component from several metals as a function of emergent proton energy.

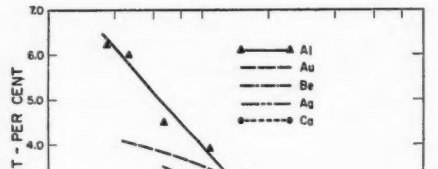


FIG. 8. H^- component from several metals as a function of emergent proton energy.

DISCUSSION OF RESULTS

1. The above data indicate that contrary to earlier investigations there are variations in the proton equilibrium charge ratios from metals. It appears likely that the previous results were due to layers of foreign material on the foil masking the ratio from the metal. The large discrepancies between observers could then be explained by the use of different pump oils, methods of trapping, and other contaminants.

2. In the present experiment only the ratios of the components of the emergent beam were measured and not the capture (σ_c) and loss (σ_l) cross sections. It can be shown that if equilibrium has been obtained and the negative component is negligible,

$$f_+/f_0=\sigma_l/\sigma_c,$$

where f_+ and f_0 are the fractions of the beam which are positive and neutral respectively.

It has been predicted[8] that the loss and capture cross sections should be proportional to some power (n) of the particle velocity and dependent on the atomic number of the material when the velocity is much larger than $v_0=e^2/\hbar$, the velocity of the hydrogen orbital electron. A summary for several cases is presented in Table II.

Applying these concepts to the present experiment, the measured ratio will take the form

$$f_+/f_0=\sigma_l/\sigma_c=kE^m,$$

where E is the energy of the particle. For light elements m should take values between 3 and 6 and for heavy elements about 3. In Fig. 9 the calculated values of m as a function of energy are shown.

In general the experimental results are in agreement with the theory. For the light elements, Be and Al, the values of m for high velocities are large (~4–5) and for the heavier elements, Ag and Au, m is almost constant over the velocity range examined and has a value of ~2.

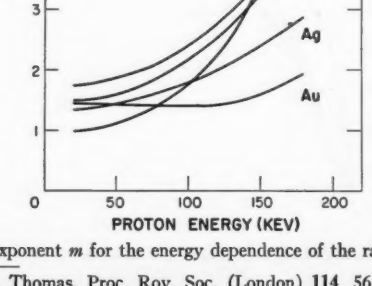


FIG. 9. Exponent m for the energy dependence of the ratio f_+/f_0.

TABLE I. Percent of positive, neutral, and negative components of a proton beam from several elements.

Element	Energy—kev 4.0	7.0	10.0	20	50	100	150	190
Dirt								
H^+	14.5%	19.7%	25.0%	37.6%	57.0%	79.9%	91.2%	96.0%
H^0	81.0	77.3	72.7	61.2	42.5	20.0	8.8	4.0
H^-	4.5	3.0	2.3	1.25	0.46	0.10	0.02	···
Aluminum								
H^+	8.2	11.2	14.2	25.1	64.0	87.9	95.5	98.2
H^0	85.6	84.0	81.8	72.5	35.6	12.0	4.5	1.8
H^-	6.2	4.8	4.0	2.4	0.37	0.06	0.04	···
Beryllium								
H^+	···	13.7	17.6	31.0	65.6	86.6	95.2	97.8
H^0	···	82.8	79.3	67.1	34.1	13.4	4.8	2.2
H^-	···	3.5	3.1	1.9	0.27	···	···	···
Calcium								
H^+	···	17.5	30.4	53.3	73.7	88.3	95.0	98.5
H^0	···	80.1	68.3	46.2	26.2	11.7	5.0	1.5
H^-	···	2.4	1.3	0.48	0.13	···	···	···
Gold								
H^+	10.2	13.2	16.5	26.7	59.5	80.4	87.6	91.5
H^0	85.6	83.0	80.0	70.5	40.1	19.5	13.4	8.5
H^-	4.2	3.8	3.5	2.8	0.4	0.09	0.05	···
Silver								
H^+	···	15.9	19.3	30.1	58.3	81.5	91.2	94.5
H^0	···	80.8	78.0	68.2	41.4	18.4	8.8	5.5
H^-	···	3.3	2.8	1.7	0.27	0.07	0.02	···
SiO								
H^+	···	20.0	24.7	36.8	59.8	83.9	92.4	···
H^0	···	77.1	73.0	61.9	39.8	16.0	7.6	···
H^-	···	2.9	2.3	1.3	0.41	0.06	0.04	···

3. The negative components decrease rapidly with increasing energy and at 150 kev are less than 0.1 percent of the total emergent beam. For the metals examined, those which have the largest neutral component at low energies (Al and Au) also have the largest negative components. In Fig. 10 is shown the ratio f_-/f_0 as a function of energy, where f_- is the fraction of the beam which is negatively charged. This ratio f_-/f_0 is also equal to the ratio of the capture and loss cross sections of the second electron if double electron loss and capture are neglected. The slopes of these curves for Al, Be, Ag, and Au have the values -0.3 ± 0.2 at low energy and -1.8 ± 0.3 at high energies. The slope

[8] L. H. Thomas, Proc. Roy. Soc. (London) **114**, 561 (1927); H. C. Brinkman and H. A. Kramers, Proc. Acad. Sci. Amsterdam **33**, 973 (1930); N. Bohr, Kgl. Danske Videnskab. Selskab, Mat.-fys. Medd. **18**, No. 8, 144 (1948).

for Ca is constant in this energy range and has a value of -0.9 ± 0.2.

Freshly evaporated Al gave the largest negative component which was 6 percent at 3 kev. At this energy, with a beam of 5 μa (limited by the foil), it is then possible to obtain a negative hydrogen beam of about 0.3 μa.

4. As the hydrogen particles leave the metal, the effect of the surface layer on the charge ratios is unknown. Oliphant and Moon,[9] Cobas and Lamb,[10] Varnerin,[11] and others have considered a similar problem for the scattering of helium ions from surfaces and the creation of the helium metastable state. It might then be expected that the charge ratios inside the metal are changed by some process dependent on a characteristic of the surface, e.g., the work function. In Fig. 11, are plotted the H^+ components against the work function of the metals at several energies. A surprisingly good correlation is found for energies above 100 kev if the value of 4.41 ev[12] be taken as the work function for evaporated Ag.

DISCUSSION OF ERRORS

As the charge ratios are determined by the last few atomic layers the results will be meaningful only if fresh clean surfaces have been used. It is not known absolutely whether, in the few seconds ($\sim$16) between an evaporation and a ratio determination, changes have occurred. The results of several tests (e.g., increasing the vacuum pressure, varying the beam current and layer thickness) indicate that if changes have taken place in this time interval they are small (<2 percent). As stated above, the calcium ratios are the most inaccurate due to the great chemical reactivity of calcium and the resulting difficulty in extrapolating to zero time.

TABLE II. Exponents for the velocity dependence of $\sigma_l \propto v^{n_l}$ and $\sigma_c \propto v^{n_c}$.

Elements	n_l	n_c
Light	-2	-5.5 to -12
Intermediate	-1	
Heavy	0	-6

Random errors are small; beam component determinations, detector sensitivity, and thickness of the foil together contribute about 3 percent to the error.

SURFACE EFFECTS

The mechanism proposed by Oliphant, Moon *et al.*,[9–11] for the ionization and neutralization of the helium atom in the field of the surface of a metal may be applicable to protons. The magnitude of such a surface effect if it occurs might be measured by keeping the charge ratio in the interior of the metal constant, i.e., the emergent beam energy constant, and varying the time the particle remains in the field of the surface. This can be done by varying the inclination of the foil to the beam while making appropriate adjustments to the beam energy.

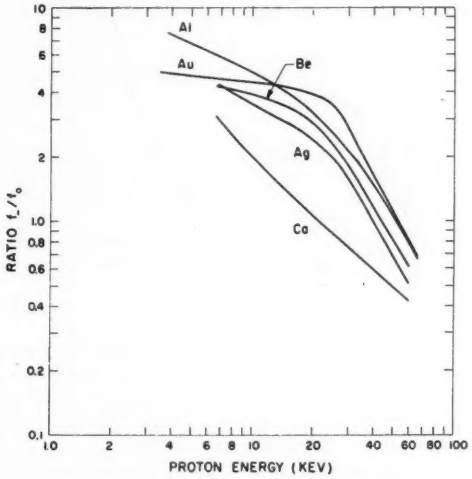


FIG. 10. Ratio f_-/f_0 as a function of energy.

The region of the surface which is of interest is small: $\sim$3 A thick. Thus to observe any effect as proposed the surface must be smooth and as free as possible from crystal structures. An amorphous surface as SiO should be ideal.

In Fig. 12, the energy loss of a 50-kev incident beam of protons passing through an inclinable foil is compared with a calculated curve assuming a $\Delta E \propto \sec\theta$ dependence and correcting for the variation of $\Delta E/\Delta x$ with energy. The good agreement ($\sim$3 percent) indicates that foils made by such an evaporation technique are uniform and that the energy lost by the beam in a foil can be calculated from the angle of inclination and the

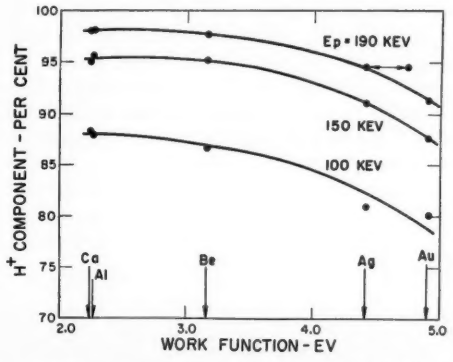


FIG. 11. H^+ components as a function of the work function of the metals at several energies.

[9] M. L. E. Oliphant and P. B. Moon, Proc. Roy. Soc. (London) A127, 388 (1930).
[10] A. Cobas and W. E. Lamb, Phys. Rev. 65, 327 (1944).
[11] L. J. Varnerin, Jr., Phys. Rev. 91, 859 (1953).
[12] L. L. Blackmer and H. E. Farnsworth, Phys. Rev. 77, 826 (1950).

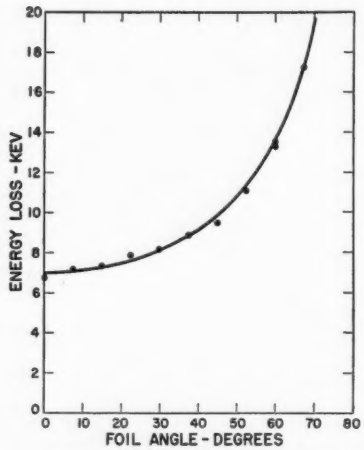


FIG. 12. Energy loss of an incident 50-kev deuteron beam in a foil as a function of the angle of the foil.

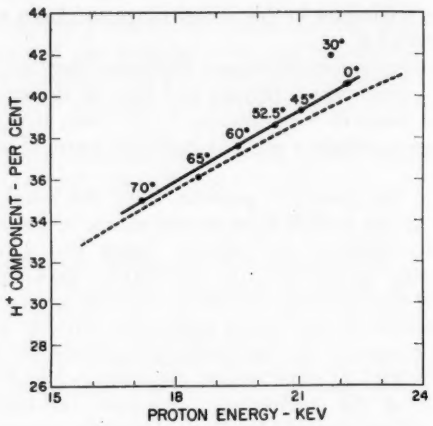


FIG. 13. Variation of the H^+ component as a function of the angle of the foil. (The dotted line is the H^+ component from a foil normal to the beam direction in which the emergent beam energy was varied.)

measured energy loss in the foil when the beam is normal to the foil.

When the H^+ component is measured from a dirty aluminum foil at different emergent energies by keeping the incident energy (50 kev) constant and varying the angle of inclination, the results shown in Fig. 13 are obtained. The dotted curve is the H^+ component from a similar foil normal to the beam direction as the emergent energy is varied by changing the incident beam energy on the foil. No effect within the experimental error (~2 percent) is found when the foil is inclined. This negative result is not conclusive in that it is known that impurities tend to condense preferentially onto a surface. The resulting rough surface would not be expected to show an effect when the foil angle is changed.

In order to obtain as flat a surface layer as possible, fresh layers of SiO and Au have been evaporated normally onto an inclinable foil during the experiment. If the beam were left normal to the foil the charge ratios would have remained relatively constant (after the initial charge following the evaporation) except for the slow growth of foreign material onto the surface as shown in Fig. 4(b). However, when the foil angle is changed the charge ratios varied with time in as yet an unpredictable manner. Though this variation in the charge ratios was small, e.g., at 50-kev incident proton energy a 6 percent change in the H^+ component has been observed, the effect was such that it overshadowed any systematic variation with angle.

CONCLUSIONS

From the evidence obtained in this experiment the following conclusions are drawn.

1. The charge ratios of a proton beam emergent from a metal surface in the energy range 3–200 kev are determined by the last few atomic layers of the material and the emergent energy of the beam.

2. The charge ratios from Ca, Al, Be, Ag, and Au are different in this energy range.

3. Above a proton energy of 100 kev there is a correlation between the charge ratios and the work function of the metal.

4. When the surface of a dirty foil is inclined at different angles to the beam direction, the charge ratios are the same as with a normal foil at the same emergent energy. Small unpredictable changes (~6 percent) are found with fresh SiO and Au when the foil is inclined.

ACKNOWLEDGMENT

The author wishes to thank James L. Tuck for many helpful discussions and to express his appreciation to R. E. Dunaway for aid in performing the experiment.

Magnetic Scattering of Slow Neutrons from O_2 Gas*

W. H. KLEINER
Lincoln Laboratory, Massachusetts Institute of Technology, Lexington, Massachusetts
(Received August 9, 1954)

A general expression is derived for magnetic scattering of slow neutrons from oxygen gas which takes account of molecular rotation. The scattering derived on the basis of Meckler's ground electronic wave function for O_2 is discussed. The validity of the static or semiclassical approximation is considered. Values are given for the total magnetic cross section in laboratory coordinates based on Meckler's wave function. The magnetic scattering does not appear to be sufficiently sensitive to changes in the wave function to distinguish in practice between different reasonable approximate wave functions.

I. INTRODUCTION

THE magnetic scattering of slow neutrons from paramagnetic O_2 has been considered as an experimental means of obtaining information about the (valence) electron distribution of the molecule.[1,2] Since the nuclear and magnetic scattering are independent, the experimental magnetic cross section may be obtained by subtracting the calculated nuclear cross section from the measured cross section. Halpern and Appleton have calculated the total nuclear cross section. In this paper a theoretical examination is made of the magnetic scattering, and the total magnetic cross section is calculated.

One might hope with either x-ray diffraction or neutron magnetic diffraction to obtain an electron distribution for the molecule from the Fourier transform of a scattering amplitude. The electron distribution measured is, of course, different for the two types of diffraction. In both cases one is immediately disappointed by the loss of information caused by molecular thermal rotational and translational motion. Neutrons of a monochromatic beam incident on a sample of gas see molecules with a variety of velocities, angular velocities, positions, and orientations (classically speaking). Assuming negligible correlation between molecules, the resultant scattering is given by averaging the cross section of one molecule with respect to these parameters. The situation is similar for an x-ray beam incident on the gas. An important distinction arises, however, from the difference in velocities of the incident particles. The speed of light is so great relative to molecular velocities and angular velocities, that a molecule appears to a photon to be stationary both in position and orientation as the photon passes by. Not so for a neutron of wavelength 1 angstrom whose speed is only 1.32×10^{-5} that of light. This added complication in the neutron case may be viewed as follows.

The ratio α of the time of transit of a neutron across a molecule to the period of molecular rotation gives a measure of how much the molecule rotates during the neutron's passage. By use of the equi-partition theorem for rotational motion, α is found to be about $\frac{1}{4}$ for O_2 gas at room temperature and incident neutrons of wavelength $\lambda=5$ A, a typical value in the range ($1\lesssim\lambda\lesssim20$ A) of experimental interest. This value of α indicates that the static (or semiclassical) approximation commonly used in x-ray scattering theory for treating the effects of molecular rotation may not be applicable to slow neutron scattering. The static approximation has been found satisfactory[3,4] under some circumstances for the nuclear scattering of slow neutrons from molecules. Here we are concerned with its validity for the magnetic scattering. The effect of translational motion of the molecules may be viewed in a similar way. The ratio of the root-mean-square molecular speed to the neutron speed gives a measure of the neutron-molecule relative velocities; for room temperature this ratio is 0.12λ(A). The translational motion of the molecules may thus also be expected to have an important bearing on the cross section.

The above qualitative discussion is from a classical viewpoint. Consider now the wave functions describing the molecule required for the quantum-mechanical treatment. We shall assume the molecule is in its ground electronic and vibrational levels. (For the gas at room temperature, the Boltzmann factor for the first excited vibrational level is 4×10^{-4}.) We neglect the zero-point vibration. The rotational levels are given by $1.44K(K+1)$ cm^{-1}, only odd-K levels being occupied because of the Bose statistics of the nuclei. The p-type tripling[5,6] of these levels arising from the spin-spin and spin-orbit terms in the Hamiltonian and the coupling of the spin with the rotational motion will be neglected. With this assumption the electronic and rotational motion become independent (except, of course, insofar as the electrons rotate with the molecule as a whole), and the electron spin of the molecule becomes free in orientation. Some of the above assumptions are considered again in Sec. IV. The molecular wave function may then be expressed

* The research in this document was supported jointly by the Army, Navy and Air Force under contract with the Massachusetts Institute of Technology.
[1] O. Halpern and G. L. Appleton, Phys. Rev. **90**, 869 (1953).
[2] H. Palevsky and R. M. Eisberg, Phys. Rev. **95**, 659 (1954).
[3] N. Z. Alcock and D. G. Hurst, Phys. Rev. **83**, 1100 (1951).
[4] J. A. Spiers, National Research Council (Canada) Report CRT-417 (April, 1949). (N.R.C. No. 1940) (unpublished).
[5] R. Schlapp, Phys. Rev. **51**, 342 (1937).
[6] J. H. Van Vleck, Revs. Modern Phys. **23**, 213 (1951).

as a product of an electronic, a vibrational, and a rotational wave function, the first determined in the Born-Oppenheimer approximation for equilibrium internuclear distance $R=1.21$ A, the second represented by a δ function of the internuclear distance, and the third by normalized spherical harmonics with the internuclear orientation as argument. The electronic wave function is a $^3\Sigma_g^-$, and will be considered in more detail in Sec. III.

Based on wave functions of this type we derive in Sec. II a general expression for the Born approximation cross section of O_2 in center-of-mass coordinates for the magnetic scattering, and also the corresponding cross section formula in the static approximation. In Sec. III we derive from Meckler's ground level electronic wave function[7] for O_2 a simple expression for the spin density in terms of which the scattering is conveniently described. In Sec. IV use of the static approximation is justified for the magnetic scattering, and the insensitivity of the scattering to the form of the $2p\pi_\pm$ orbital of the wave function is discussed. The total magnetic cross section is evaluated in Sec. V.

II. MAGNETIC SCATTERING CROSS-SECTION FORMULAS

A parallel beam of unpolarized slow neutrons of definite wavelength is assumed to be incident on a sample of O_2 gas in thermal equilibrium. No correlation between the molecules being assumed, the cross section for the gas is the sum of the cross sections for the individual molecules. Double scattering is negligible. The cross section in center-of-mass coordinates for the magnetic scattering of an unpolarized neutron beam may be expressed in terms of the partial cross sections per initial state of the molecule,

$$\sigma_{ba}(\boldsymbol{\kappa})=(k'/k)d_a^{-1}\sum(\nu_b',\nu_a)(\mathbf{I}-\mathbf{ee}) : (\nu_b'|\mathbf{P}|\nu_a)^*(\nu_b'|\mathbf{P}|\nu_a), \quad (1)$$

corresponding to incident neutrons of wave vector $\mathbf{k}$ scattered from a molecule initially in an energy level E_a. After scattering, the molecule is left with an energy E_b and the neutron with wave vector $\mathbf{k}'$, with k' determined by conservation of energy: $\hbar^2(k'^2-k^2)/2\mathfrak{m}=E_a-E_b$, where $\mathfrak{m}$ is the neutron-molecule reduced mass. The momentum transferred to the molecule is $\hbar(\mathbf{k}-\mathbf{k}')=\hbar\boldsymbol{\kappa}=\hbar\kappa\mathbf{e}$. The summation is over the degenerate states, denoted by ν_a and ν_b', of the energy levels E_a and E_b; d_a is the degeneracy of E_a. σ_{ba} is given in (1) in units of $[(\mathfrak{m}/m_p)(g_n e^2/2mc^2)]^2$; m_p is the proton mass, g_n the neutron g factor. $\mathbf{I}$ is the unit dyadic. The magnetic scattering operator $\mathbf{P}$ in (1) is the sum of a spin part $\mathbf{P}^{(S)}$ and an orbital part[8] $\mathbf{P}^{(L)}$:

$$\mathbf{P}^{(S)}(\boldsymbol{\kappa})=\sum_j \exp(i\boldsymbol{\kappa}\cdot\mathbf{r}_j)\mathbf{s}_j, \quad (2a)$$

$$\mathbf{P}^{(L)}(\boldsymbol{\kappa})=\sum_j \tfrac{1}{4}\{E(\boldsymbol{\kappa}\cdot\mathbf{r}_j)\mathbf{l}_j+\mathbf{l}_j E(\boldsymbol{\kappa}\cdot\mathbf{r}_j)\}, \quad (2b)$$

where

$$E(\boldsymbol{\kappa}\cdot\mathbf{r})=2\int_0^1 \lambda\exp(i\lambda\boldsymbol{\kappa}\cdot\mathbf{r})d\lambda \quad (3)$$

$$\to 1 \quad \text{as} \quad \kappa\to 0.$$

The position, spin, and orbital angular momentum operators of the jth electron are denoted by $\mathbf{r}_j$, $\mathbf{s}_j$, and $\mathbf{l}_j$. The cross section per initial state of a molecule with energy E_a is obtained by summing σ_{ba} over all levels E_b consistent with conservation of energy. This cross section is then averaged over all initial states weighted by the Boltzmann factor.

A typical molecular wave function, discussed in Sec. I, will be denoted by $Y_J^M(\mathbf{u})\psi_e^{M_S}(X_1',X_2',\cdots,X_{16}')$; the rotational wave function[9] Y_J^M is a normalized spherical harmonic[10] referred to a frame of reference fixed in space; the unit vector $\mathbf{u}$ defines the orientation of the internuclear axis. $\psi_e^{M_S}$ is the $^3\Sigma_g^-$ ground electronic wave function with spin magnetic quantum number M_S; it is convenient to quantize the spin $\mathbf{S}$ with respect to the space reference frame; X_j' denotes the position and spin coordinates of the jth electron, the prime indicating that the space coordinates are referred to a reference frame fixed in the molecule. The degeneracy indices ν of (1) refer to M and M_S.

The components of the vectors in (2) are ordinarily defined with respect to the space frame. To evaluate the matrix elements of (1) it is convenient to express the $\mathbf{r}_j$ and $\mathbf{l}_j$ with respect to the molecular frame. Using the invariance of a scalar product under rotation of coordinate reference frame: $\boldsymbol{\kappa}\cdot\mathbf{r}=\boldsymbol{\kappa}'\cdot\mathbf{r}'$ and $\mathbf{e}\cdot\mathbf{P}^{(L)}=\mathbf{e}'\cdot\mathbf{P}^{(L)'}$ we may express the matrix element $(J'M'M_S'|\mathbf{P}^{(L)}(\boldsymbol{\kappa})|\times JMM_S)$ in terms of

$$\mathbf{f}^{(L)'}(\boldsymbol{\kappa}')=\delta(M_S',M_S)\int\psi_e^{M_S*}(X') \times\sum_j\tfrac{1}{4}\{E(\boldsymbol{\kappa}'\cdot\mathbf{r}_j')\mathbf{l}_j'+\mathbf{l}_j'E(\boldsymbol{\kappa}'\cdot\mathbf{r}_j')\} \times\psi_e^{M_S}(X')dX_e', \quad (4)$$

where the integral is over the electron coordinates and includes summation over spins; $\mathbf{f}^{(L)'}$ is independent of M_S. $\mathbf{f}^{(L)'}$ vanishes for a Σ state, and as a consequence the orbital interaction leads to no scattering. It is well known that the average orbital angular momentum vanishes for an orbitally nondegenerate level. The usual proof, based on time reversal symmetry, may be generalized to show that $\mathbf{f}^{(L)'}$ vanishes: denoting the Hermitian conjugate by †, we have

$$\mathbf{l}^\dagger=\mathbf{l}=-\mathbf{l}^*; \quad [\exp(i\boldsymbol{\kappa}\cdot\mathbf{r})]^\dagger =\exp(-i\boldsymbol{\kappa}\cdot\mathbf{r})=[\exp(i\boldsymbol{\kappa}\cdot\mathbf{r})]^*,$$

[7] A. Meckler, J. Chem. Phys. **21**, 1750 (1953).

[8] The convenient form of the orbital operator given here is due to G. T. Trammel, Phys. Rev. **92**, 1387 (1953).

[9] The total orbital angular momentum quantum number is commonly denoted by K. See, e.g., reference 6.

[10] E. U. Condon and G. H. Shortley, *The Theory of Atomic Spectra* (Cambridge University Press, London, 1951).

and hence

$$\mathbf{P}^{(L)}(\boldsymbol{\kappa})^\dagger=\mathbf{P}^{(L)}(-\boldsymbol{\kappa})=-\mathbf{P}^{(L)}(\boldsymbol{\kappa})^*.$$

Upon dropping the primes for simplicity in (4), $\mathbf{f}^{(L)\prime}(\boldsymbol{\kappa}')$ becomes

$$\mathbf{f}^{(L)}(\boldsymbol{\kappa})=\delta(M_S',M_S)\int\psi_e^*\mathbf{P}^{(L)}(\boldsymbol{\kappa})\psi_e dX_e$$
$$=-\delta(M_S',M_S)\int\psi_e\mathbf{P}^{(L)}(\boldsymbol{\kappa})\psi_e^* dX_e,$$

and, since ψ_e can be chosen real, it follows that $\mathbf{f}^{(L)}$ vanishes.

The matrix elements of the spin scattering operator may be treated similarly in terms of the integral

$$\int\psi_e^{M_S'*}\textstyle\sum_j\exp(i\boldsymbol{\kappa}'\cdot\mathbf{r}_j')\mathbf{s}_j\psi_e^{M_S}dX_e'.$$

The operator $\mathbf{P}^{(S)}$ obeys vector type commutation relations[10] with respect to $\mathbf{S}=\sum_j\mathbf{s}_j$; i.e., $[P_x^{(S)},S_y]=i\hbar P_z$, etc. Consequently, the integral may be reduced to

$$(SM_S'|\mathbf{S}|SM_S)\left\{\int\psi_e^{S*}\textstyle\sum_j\exp(i\boldsymbol{\kappa}'\cdot\mathbf{r}_j')s_{jz}\psi_e^S dX_e'/S\right\}.$$

Equation (1) may now be simplified by introducing the latter and using the relation

$$\sum_{M_S',M_S}(\mathbf{I}-\mathbf{ee}):(M_S'|\mathbf{S}|M_S)^*(M_S'|\mathbf{S}|M_S)$$
$$=\tfrac{2}{3}S(S+1)(2S+1).$$

Thus,

$$\sigma_{J'J}(\boldsymbol{\kappa})=[(2S+1)(2J+1)]^{-1}(k'/k)$$
$$\times\sum_{M',M,M_S',M_S}(\mathbf{I}-\mathbf{ee})$$
$$:(J'M'M_S'|\mathbf{P}^{(S)}(\boldsymbol{\kappa})|JMM_S)^*$$
$$\times(J'M'M_S'|\mathbf{P}^{(S)}(\boldsymbol{\kappa})|JMM_S)$$
$$=\tfrac{2}{3}S(S+1)(2J+1)^{-1}(k'/k)$$
$$\times\sum_{M',M}|(J'M'|f(\boldsymbol{\kappa}')|JM)|^2,\quad(5)$$

where

$$f(\boldsymbol{\kappa}')\equiv\int\psi_e^{S*}\textstyle\sum_j\exp(i\boldsymbol{\kappa}'\cdot\mathbf{r}_j')s_{jz}\psi_e^S dX_e'/S\quad(6a)$$
$$\to1\quad\text{as}\quad\boldsymbol{\kappa}\to0.$$

The integral of $\psi_e^{S*}\exp(i\boldsymbol{\kappa}'\cdot\mathbf{r}_j')s_{jz}\psi_e^S$ over all electron coordinates (including spin) except $\mathbf{r}_j'$ is independent of j, because of the antisymmetry of ψ_e in the electron coordinates. This allows $f(\boldsymbol{\kappa}')$ to be expressed as a Fourier transform,

$$f(\boldsymbol{\kappa}')=\int\exp(i\boldsymbol{\kappa}'\cdot\mathbf{r}')\rho(\mathbf{r}')d\tau',\quad(6b)$$

of a normalized spin density

$$\rho(\mathbf{r}')\equiv N\int\psi_e^{S*}(X',X_2',X_3'\cdots,X_{16}')$$
$$\times s_z\psi_e^S(X',X_2',X_3'\cdots,X_{16}')d\sigma dX_2'\cdots,dX_{16}'/S,\quad(7a)$$

$d\sigma$ denoting summation over spin; $\int\rho(\mathbf{r}')d\tau'=1$. The quantity ρ, and consequently f, has the symmetry of the molecule. In evaluating the matrix elements of (5) it is convenient to expand $f(\boldsymbol{\kappa}')$ in spherical harmonics:

$$f(\boldsymbol{\kappa}')=\sum_{l=0}^{\infty}(2l+1)i^lA_l(\kappa)P_l(\mathbf{e}'\cdot\mathbf{u}'),\quad(6c)$$

where $A_l(\kappa)$ is given by[11]

$$A_l(\kappa)\equiv\int P_l(\mathbf{e}'\cdot\mathbf{u}')f(\boldsymbol{\kappa}')d\mathbf{e}'/4\pi i^l\quad(8a)$$

$$=\int j_l(\kappa r')P_l(\mathbf{u}'\cdot\mathbf{r}'/r')\rho(\mathbf{r}')d\tau';\quad(8b)$$

or, if the spin density (7a) is expanded in spherical harmonics

$$\rho(\mathbf{r}')=\sum_{l=0}^{\infty}[(2l+1)/4\pi]\rho_l(r')P_l(\mathbf{u}'\cdot\mathbf{r}'/r'),\quad(7b)$$

by

$$A_l(\kappa)=\int_0^\infty j_l(\kappa r')\rho_l(r')r'^2dr'.\quad(8c)$$

Let $\mathfrak{R}^{-1}$ denote a rotation which sends the space frame into the molecular frame. We choose the z axis of the molecular frame along $\mathbf{u}$ and the z axis of the space frame along $\mathbf{e}$; the coordinates of $\mathbf{e}$ and $\mathbf{u}'$ are then the same, namely (0,0,1) in Cartesian coordinates. Symbolically the coordinates of $\mathbf{u}$ may be described by $\mathbf{u}=\mathfrak{R}^{-1}\mathbf{e}$ and the coordinates of $\boldsymbol{\kappa}'=\kappa\mathbf{e}'$ by $\mathbf{e}'=\mathfrak{R}\mathbf{u}'$, or, since the coordinates of $\mathbf{e}$ and $\mathbf{u}'$ are identical, by $\mathbf{e}'=\mathfrak{R}\mathbf{e}$. To facilitate the evaluation of the matrix elements

$$(J'M'|f(\boldsymbol{\kappa}')|JM)$$
$$=\int Y_J^{M'*}(\mathfrak{R}^{-1}\mathbf{e})f(\mathfrak{R}\boldsymbol{\kappa})Y_J^M(\mathfrak{R}^{-1}\mathbf{e})d\mathbf{u}\quad(9)$$

of (5), in which the integration is over all solid angle, it is desirable to have the same argument appearing in each factor of the integrand. The azimuthal direction of the x- (or y-) axis of the molecular frame may be chosen arbitrarily because of the cylindrical symmetry of the molecule. We make use of this freedom in defining $\mathfrak{R}^{-1}$ as a rotation about an axis perpendicular to the plane of $\mathbf{e}$ and $\mathbf{u}$. Then, if θ, ϕ are the spherical coordinates of $\mathbf{u}=\mathfrak{R}^{-1}\mathbf{e}$, $\theta,\phi\pm\pi$ are the spherical coordinates of $\mathbf{e}'=\mathfrak{R}\mathbf{e}$.

[11] The P_l are Legendre polynomials; the j_l are spherical Bessel functions; $d\mathbf{e}'$ denotes the element of solid angle.

From the properties of the spherical harmonics it follows that

$$Y_J^M(\mathfrak{R}^{-1}\mathbf{e})=Y_J^M(\theta,\phi)$$
$$=(-)^M Y_J^M(\theta,\phi\pm\pi)=(-)^M Y_J^M(\mathfrak{R}\mathbf{e}),$$

so that the arguments in the factors of the integrand of (9) may be made the same. The cross section $\sigma_{J'J}$ will depend only on the magnitude of $\boldsymbol{\kappa}$. The average with respect to molecular orientations associated with $(J'M'|f|JM)$ is equivalent to an average over the orientations of $\boldsymbol{\kappa}$.

The summation in (5) may be carried out with the aid of the spherical harmonic addition theorem

$$\sum_M Y_J^{M*}(\mathbf{u}_1)Y_J^M(\mathbf{u}_2)=(2J+1)P_J(\mathbf{u}_1\cdot\mathbf{u}_2)/4\pi \quad (10)$$

to yield

$$\sigma_{J'J}(\kappa)=\tfrac{2}{3}S(S+1)(2J'+1)(k'/k)$$
$$\times\int P_{J'}(\mathbf{u}_1\cdot\mathbf{u}_2)P_J(\mathbf{u}_1\cdot\mathbf{u}_2)f^*(\kappa\mathbf{u}_1)$$
$$\times f(\kappa\mathbf{u}_2)d\mathbf{u}_1d\mathbf{u}_2/(4\pi)^2. \quad (11)$$

Since $P_J(\mathbf{u}_1\cdot\mathbf{u}_2)$ is invariant under simultaneous rotation of $\mathbf{u}_1$ and $\mathbf{u}_2$, it follows from the orthogonality theorem for the irreducible representations of the rotation group that $\int P_{J'}(\mathbf{u}_1\cdot\mathbf{u}_2)P_J(\mathbf{u}_1\cdot\mathbf{u}_2)Y_{l'}^{m'*}(\mathbf{u}_1\cdot\mathbf{e})$ $Y_l^m(\mathbf{u}_2\cdot\mathbf{e})d\mathbf{u}_1d\mathbf{u}_2$ vanishes unless $l'=l$ and $m'=m$, and is independent of m for given l. This fact, together with the addition theorem (10), allows one to evaluate in terms of [12,13]

$$\int_{-1}^{1}P_{J'}(\mu)P_J(\mu)P_l(\mu)d\mu=[2/(2l+1)](J'J00|J'Jl0)^2$$

the type of integral obtained when (6c) is substituted into (11). The result is

$$\sigma_{J'J}(\kappa)=\tfrac{2}{3}S(S+1)(2J'+1)(k'/k)$$
$$\times\sum_{l=0}^{\infty}(J'J00|J'Jl0)^2A_l^2(\kappa). \quad (12)$$

Consider now two limiting cases for the cross section corresponding to large and small values of the ratio α (mentioned in Sec. I) of the transit time of the neutron to the period of molecular rotation. The case $\alpha\gg1$ is academic, but instructive: very cold neutrons are incident on a "gas" at very low temperature, and the molecules are supposed to have no translational motion. From a classical viewpoint a neutron sees many orientations of a molecule during transit.

$$f(\boldsymbol{\kappa}')=f(\mathfrak{R}\boldsymbol{\kappa})=\int\exp(i\boldsymbol{\kappa}\cdot\mathfrak{R}^{-1}\mathbf{r}')\rho(\mathbf{r}')d\tau'$$
$$=\int\exp(i\boldsymbol{\kappa}\cdot\mathbf{r}')\rho(\mathfrak{R}\mathbf{r}')d\tau'=\int\exp(i\boldsymbol{\kappa}\cdot\mathbf{r})\rho(\mathfrak{R}\mathbf{r})d\tau$$

may be regarded as a scattering amplitude for a stationary molecule, so that $|f(\boldsymbol{\kappa}')|^2$ gives the cross section (to within a constant factor). For $\alpha\gg1$, then, the amplitude should correspond to an f with the spin density ρ averaged uniformly over all orientations of the molecule, which is equivalent to averaging $f(\boldsymbol{\kappa})$ over all orientations of $\boldsymbol{\kappa}$; i.e., symbolically, the amplitude should be $\langle f(\boldsymbol{\kappa})\rangle$, and the cross section $|\langle f\rangle|^2$. The quantum analog has the molecule initially in its lowest rotational level, and an incident neutron has insufficient energy for inelastic scattering; i.e., $J'=J=0$. In agreement with the classical picture, we obtain

$$\sigma_{00}(\kappa)=\tfrac{2}{3}S(S+1)|\langle f\rangle|^2=\tfrac{2}{3}S(S+1)A_0^2(\kappa). \quad (13)$$

For $\alpha\ll1$, the *static* or *semiclassical* limiting case, each neutron sees an instantaneous picture of the molecule corresponding to a cross section $|f(\boldsymbol{\kappa}')|^2$, but the random orientations of the molecule require this to be averaged so that for the gas the cross section is $\langle|f|^2\rangle$. In the quantum analog one envisages such large values of k that k'/k may be replaced by unity in $\sigma_{J'J}$. Then, with the aid of the closure property for spherical

TABLE I. Determinants of the symmetry states. The normalized determinants $A, \cdots, \phi_f$ with $M_S=1$ are defined by the molecular orbitals on the principal diagonal; these consist of filled $1s\sigma$ and $2s\sigma$ orbitals accounting for eight of the sixteen electrons, the remaining orbitals being specified by the table, a + or − entry indicating that the space orbital is occupied by an electron with + or − spin. The space orbitals are labeled both in standard notation and in Meckler's notation. Meckler lists the nine $^3\Sigma_g^-$ symmetry states as eigenfunctions of S_z with eigenvalue $M_S=0$; these yield the ϕ_μ when operated on by $(S_x+iS_y)/\sqrt{2}$.

	$2p\sigma_g$ ϕ_0	$2p\pi_u^+$ ϕ_+	$2p\pi_u^-$ ϕ_-	$2p\pi_g^+$ χ_+	$2p\pi_g^-$ χ_-	$2p\sigma_u$ χ_0
A	+ −	+	+	+	−	+ −
B	+ −	+	−	+	+	+ −
C	+ −	+	+	−	+	+ −
D	+ −	−	+	+	+	+ −
E	+	+ −	−	+	+ −	+
F	+	+ −	+	−	+ −	+
G	+	+ −	+	+	+ −	−
H	−	+ −	+	+	+ −	+
I	+	+	+ −	+ −	−	+
J	+	−	+ −	+ −	+	+
K	+	+	+ −	+ −	+	−
L	−	+	+ −	+ −	+	+
ϕ_c	+ −	+ −	+ −	+		
ϕ_d	+ −	+	+ −	+ −	+ −	
ϕ_e		+ −	+ −	+	+	+ −
ϕ_f		+	+	+ −	+ −	+ −

[12] See, e.g., Biedenharn, Blatt, and Rose, Revs. Modern Phys. 24, 249 (1952) for a discussion of the vector addition coefficients $(j_1j_2m_1m_2|j_1j_2jm)$, and a convenient computational formula. Note that $P_{J'}P_J=\sum_{l=0}^{\infty}(J'J00|J'Jl0)^2P_l$.

[13] M. Hamermesh and J. Schwinger, Phys. Rev. 69, 145 (1946) give an analogous derivation for the neutron-nuclear scattering from deuterium.

harmonics

$$\sum_J \sum_M Y_J^{M*}(\mathbf{u}_1) Y_J^M(\mathbf{u}_2) = \sum_J (2J+1) P_J(\mathbf{u}_1 \cdot \mathbf{u}_2)/4\pi = \delta(\mathbf{u}_1 - \mathbf{u}_2),$$

the cross section per initial state for a molecule initially in level E_J becomes, from (5),

$$\sum_{J'} \sigma_{J'J}(\kappa) = \tfrac{2}{3} S(S+1) \langle |f(\kappa)|^2 \rangle, \tag{14}$$

independent of J.

III. THE SPIN DENSITY

The spin density[14] $\rho(\mathbf{r})$ of Eqs. (7) contains all the information about the ground electronic wave function ψ_e^S of the molecule relevant to the scattering problem. More generally, a spin density may be used to express in a convenient form matrix elements diagonal in S of any operator of the type[10] $\mathbf{F} = \sum_j h(\mathbf{r}_j)\mathbf{s}_j$: $(SM_S'|\mathbf{F}|SM_S) = (SM_S'|\mathbf{S}|SM_S) \int h(\mathbf{r})\rho(\mathbf{r})d\tau$. The spin scattering operator (2a) corresponds to choosing $h(\mathbf{r}) = \exp(i\boldsymbol{\kappa}\cdot\mathbf{r})$.

The normal O_2 electronic level is a $^3\Sigma_g^-$. Meckler[7] approximates ψ_e^S for this level by a linear combination of nine symmetry states of $^3\Sigma_g^-$ symmetry, each a linear combination of Slater determinants of (orthonormal) molecular orbitals, the orbitals being approximated by linear combinations of atomic orbitals. The molecular space orbitals for each determinant are chosen from the set

$$1s\sigma_g,\ 1s\sigma_u,\ 2s\sigma_g,\ 2s\sigma_u,\ 2p\sigma_g,\ 2p\pi_u^{\pm},\ 2p\pi_g^{\pm},\ 2p\sigma_u.$$

At normal internuclear distance the one-electron energies of the orbitals increase from left to right in the list. We write Meckler's ψ_e^S as

$$\psi = \sum_{\mu=a}^{i} c_\mu \phi_\mu \quad \left(\sum_{\mu=a}^{i} c_\mu^2 = 1\right), \tag{15}$$

where, in terms of the determinants specified in Table I, the symmetry states ϕ_μ are given by

$$\begin{aligned}
\phi_a &= \tfrac{1}{2}(A+B-C-D),\\
\phi_b &= \tfrac{1}{2}(B+C-A-D),\\
\phi_g &= 2^{-\frac{1}{2}}(E+G-F-H+J+K-I-L),\\
\phi_h &= 2^{-\frac{1}{2}}(F+G-E-H+I+K-J-L),\\
\phi_i &= 2^{-\frac{1}{2}}(E+F-G-H+I+J-K-L),
\end{aligned} \tag{16}$$

and the one-determinant symmetry states ϕ_c, ϕ_d, ϕ_e, and ϕ_f. The symmetry state ϕ_c, which corresponds to the molecular orbital solution, predominates at normal internuclear distance $(1-|c_c|^2 = 0.05)$. The set of orbitals for ϕ_c gives the lowest value for the sum of one-electron energies.

[14] Henceforth we drop the primes and refer all coordinates to the molecular frame.

TABLE II. Values of the C's of Eqs. (17) as a function of internuclear distance for Meckler's O_2 normal $^3\Sigma_g^-$ electronic wave function. The importance of the one-determinant molecular orbital solution ϕ_c in Meckler's wave function is indicated in the last row. Note that $\sum_i C_i = 1$.

R(a.u.)	2.0	2.5	4.0	∞
C_1	−0.0009	−0.0018	0.2212	1/4
C_2	0.0180	0.0394	0.2590	1/4
C_3	0.9816	0.9591	0.2980	1/4
C_4	0.0013	0.0033	0.2213	1/4
c_c^2	0.9700	0.9347	0.1902	1/8

In case, as here, ψ_e^S is expressed as a linear combination of determinants of orthonormal orbitals, $\rho(\mathbf{r})$ is a quadratic form in the orbitals, the cross terms (overlap terms) arising from products of determinants differing by one orbital. No overlap terms appear in ρ for Meckler's ψ_e^S, since all determinants involved differ by at least two orbitals. This follows from the fact that the $1s\sigma$ and $2s\sigma$ orbitals are always filled and no two space orbitals outside the $1s\sigma$ and $2s\sigma$ are of the same symmetry; changing the symmetry of only one space orbital would change the symmetry of ψ_e^S.

The result of inserting (15) into (7a) is, in terms of the molecular space orbitals φ,

$$\begin{aligned}
\rho(\mathbf{r}) = {} & C_1|\varphi(2p\sigma_g)|^2 + C_2|\varphi(2p\pi_u^{\pm})|^2\\
& + C_3|\varphi(2p\pi_g^{\pm})|^2 + C_4|\varphi(2p\sigma_u)|^2,
\end{aligned} \tag{17a}$$

where

$$\begin{aligned}
C_1 &= \alpha+\beta+\gamma-\delta,\\
C_2 &= \tfrac{1}{2}|c_a-c_b|^2 + c_d^2 + c_f^2 - \alpha+\beta+\gamma+\delta,\\
C_3 &= \tfrac{1}{2}|c_a+c_b|^2 + c_c^2 + c_e^2 + \alpha-\beta+\gamma+\delta,\\
C_4 &= \alpha+\beta-\gamma+\delta,\\
\alpha &= \tfrac{1}{8}(c_g-c_h+c_i)^2, \quad \beta = \tfrac{1}{8}(-c_g+c_h+c_i)^2,\\
\gamma &= \tfrac{1}{8}(c_g+c_h-c_i)^2, \quad \delta = \tfrac{1}{8}(c_g+c_h+c_i)^2.
\end{aligned} \tag{17b}$$

Note that $\sum_i C_i = 1$. The coefficients c_μ are given by Meckler as a function of internuclear distance; corresponding values of the coefficients C_i are given in Table II. At equilibrium distance the contribution of the $2p\sigma$ orbitals may be neglected, and we may write $C_2 + C_3 = 1$. With the $2p_\pm$ atomic orbitals at $-\mathbf{R}/2$ and $\mathbf{R}/2$ denoted by ϕ and ϕ', and the overlap integral $\int \phi^*\phi' d\tau$ by Δ, the spin density (17a, b) reduces to

$$\rho(\mathbf{r}) = a(|\phi|^2 + |\phi'|^2) + b2\phi^*\phi', \tag{18}$$

with

$$a = \frac{1+(2C_3-1)\Delta}{2(1-\Delta^2)}, \quad b = -\frac{(2C_3-1)+\Delta}{2(1-\Delta^2)}. \tag{18'}$$

These reduce to $a = -b = \frac{1}{2}$ for the molecular orbital solution together with $\Delta = 0$. For $C_3 = 0.97$ (interpolation for $R = 2.28$ from Table II) and reasonable values of Δ, a and $-b$ may be larger by as much as about 30 percent. (For Meckler's Gaussian, Duncanson-Coulson's hydro-

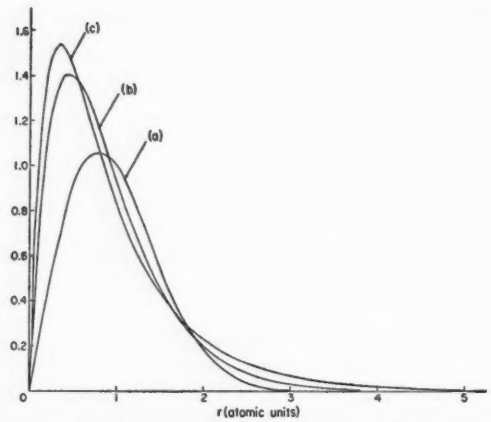


FIG. 1. $R(r)$ for normalized atomic $2p_{\pm}$ orbitals $R(r)Y_1^{\pm 1}(\theta,\phi)$ of the (a) Meckler-Gaussian (see reference 7), (b) hydrogenic (see reference 15) ($Z=4.44$), and (c) Hartree-Fock (see reference 16) types.

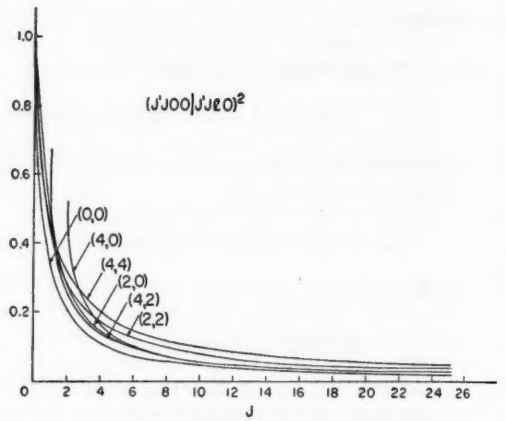


FIG. 2. A plot of $(J'J00|J'J\mathcal{l}0)^2$ *vs* J. These functions are here defined only for integral values J and J'. The smooth curves are drawn for clarity. The curves are labeled by $(l,\Delta J)$ where $\Delta J = J'-J$.

genic,[15] and Hartree-Fock[16] orbitals, Δ is 0.125, 0.158, and 0.244, respectively, at $R=2.28$ a.u.) The radial wave functions for these orbitals are plotted in Fig. 1 for comparison. If Hartree-Fock or other reasonable types of atomic orbitals were used in Meckler's calculation in place of his Gaussian-type orbital, the c_μ might differ somewhat, but it seems unlikely that the C's of (17) would change significantly.

IV. MAGNETIC CROSS SECTION IN CENTER-OF-MASS COORDINATES

From a study of computed curves of the $A_l(\kappa)$, which depend on the spin density through $f(\kappa)$, two main conclusions may be drawn. First, the error in the cross section associated with the static approximation may generally be neglected. This is fortunate, since use of the static approximation allows very great reduction in the amount of computing required to obtain the cross section. Second, the cross section is rather insensitive to the form of the radial part of the oxygen $2p$ atomic orbital. This is unfortunate in that it means the scattering does not provide a sensitive probe for the electron distribution of the molecule.

The validity of the static approximation (14) is based on the smallness of the inelastic scattering relative to the elastic scattering. $(J'J00|J'J\mathcal{l}0)$ vanishes for $J'+J+l$ odd and for $|J'-J|>l$. From the inversion symmetry of the molecule it follows that A_l vanishes for l odd, and hence only J' and J for even $|J'-J|$ are involved. (This is independently a consequence of the symmetry of the wave function under exchange of nuclei if the molecule contains only one isotope.) Some values of $(J'J00|J'J\mathcal{l}0)^2$ are plotted in Fig. 2. It follows that if $A_l=0$ for $l\neq 0$, i.e., if the spin density is spherically symmetrical, the scattering is purely elastic. In this case the static approximation is justified, since $\sigma_{J'J}=\delta_{J'J}\sigma_{JJ}$, so that replacing k'/k by unity, as required in the static approximation, has rigorously no effect on $\sigma_{J'J}$, and including contributions $\sigma_{J'J}$ to the scattering from transitions for which energy cannot be conserved, also required by the static approximation, also has no effect. The actual scattering approaches this ideal case. The $A_l(\kappa)$ derived from Meckler's ground electronic wave function (15) are plotted in Fig. 3. Denoting the maximum value of $|A_l(\kappa)|$ for given l by $|A_l|_m$, we may use the smallness of $(|A_l|_m/|A_0|_m)^2=|A_l|_m^2$ as a criterion for the validity of the static approximation; if this ratio vanishes for all $l>0$, we have the ideal case discussed above. It may be seen from Table III that this criterion is quite well satisfied. It is true that this criterion is no longer satisfactory if values of κ for which $|A_0(\kappa)|\approx|A_2|_m$ are of particular interest, e.g., in differential cross section measurements for sufficiently large scattering angles and/or small incident wave lengths; such cases, which involve smaller scattered intensities (especially relative to the nuclear scattering), are likely to be of less experimental interest in the near future.

Another indication that use of the static approximation will lead to satisfactory results comes from a comparison of $\langle|f|^2\rangle$ and $|\langle f\rangle|^2$ discussed in Sec. II. These differ appreciably only as κ approaches the tail region (see Fig. 4), and since they represent two limiting cases for the ratio α, one might expect intermediate cases not to differ significantly. We may also interpret $|\langle f\rangle|^2$ as the rigorous scattering if $A_l=0$ for $l\neq 0$, and the difference between $\langle|f|^2\rangle$ and $|\langle f\rangle|^2$ as taking approximate account of the nonvanishing of

[15] W. E. Duncanson and C. A. Coulson, Proc. Roy. Soc. Edinburgh **A62**, 37 (1944).

[16] Hartree, Hartree, and Swirles, Trans. Roy. Soc. (London) **A238**, 229 (1939). To calculate Δ we used the analytic fit of the radial function for the oxygen ^1S given by P. O. Löwdin, Phys. Rev. **90**, 120 (1953), together with formulas from C. C. J. Roothaan, J. Chem. Phys. **19**, 1445 (1951).

these A_l for $l \neq 0$, i.e., the nonsphericity of the spin density.

For the nuclear scattering one finds a partial cross section[13,8] "$\sigma_{J'J}$" similar to (12). The κ dependence of the nuclear partial cross section is obtained by replacing the electron spin density by the density of nuclear matter, i.e., δ functions at the equilibrium nuclear positions (if vibration is ignored): $f(\kappa)$ and $A_l(\kappa)$ reduce simply to $\cos(\boldsymbol{\kappa}\cdot\mathbf{R}/2)$ and $\frac{1}{2}[1+(-)^l]j_l(\kappa R/2)$. From a comparison of the A_l for the magnetic and nuclear scattering plotted in Fig. 3, it is apparent that the static approximation has greater validity in the magnetic case. There are two reasons for this. First, the spin density is widely distributed radially so that its Fourier transform falls off more rapidly with κ than in the nuclear case where the density is finite only on one spherical shell. Second, the spin density is more spherically distributed than the nuclear density, as indicated by (7b) and (8c).

The A_l were computed with a hydrogenic $2p_\pm$ atomic orbital with effective nuclear charge[15] $Z=4.44$ in (18) replacing Meckler's Gaussian orbital (but maintaining $C_3=0.97$). The A_l are very similar in form for the two types of orbitals, both behaving roughly like damped spherical Bessel functions (see Fig. 3), the damping being in agreement with the extensive radial distribution of the spin density. Table III indicates the relative positions of the main peaks for each l. The A_0's differ by at most 0.03 for the two orbitals (see Fig. 3). It appears unlikely that the cross sections determined from these two orbitals could be experimentally distinguished at present. It may be of interest that the contribution to A_0 from the b or overlap part of (18) is relatively small; it has the general shape of A_0 itself, but falls off from about -0.15 at $\kappa=0$ crossing the axis at about $\kappa R/2=3.5$ and 7.

As an experimental means of distinguishing between various reasonable types of molecular wave functions of O_2, the scattering of neutrons proves to be of little use. We see that even before the blurring effect of the translational motion of the molecules is taken into account, the magnetic scattering differs very little for the two types of $2p_\pm$ orbitals assumed. Our original assumptions for the molecular wave function disregarded the zero-point vibration (for which the rms fractional

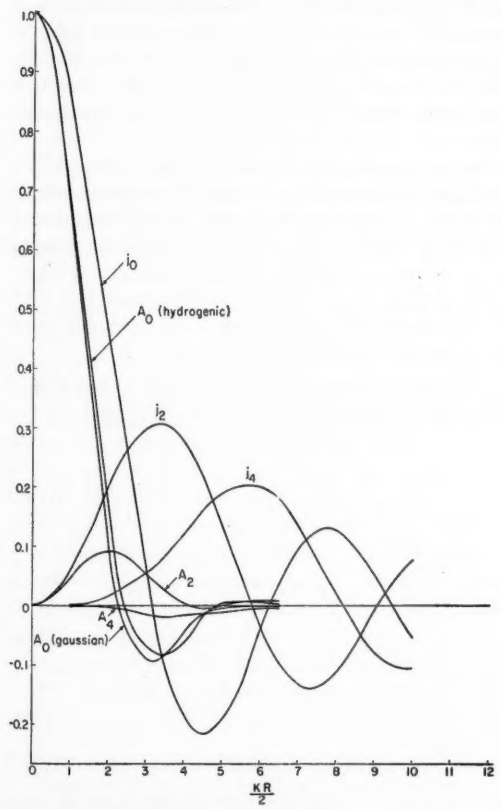


FIG. 3. The coefficients A_l, $l=0$, 2, 4, of Eqs. (8) for Meckler's ground electronic wave function *vs* $\kappa R/2$ are plotted together with j_l, $l=0$, 2, 4, the corresponding coefficients for the nuclear scattering. The small effect on the magnetic scattering of replacing the Gaussian $2p$ atomic orbital of Meckler's wave function by a hydrogenic orbital ($Z=4.44$) may be seen by comparing the A_0's plotted for the two orbitals.

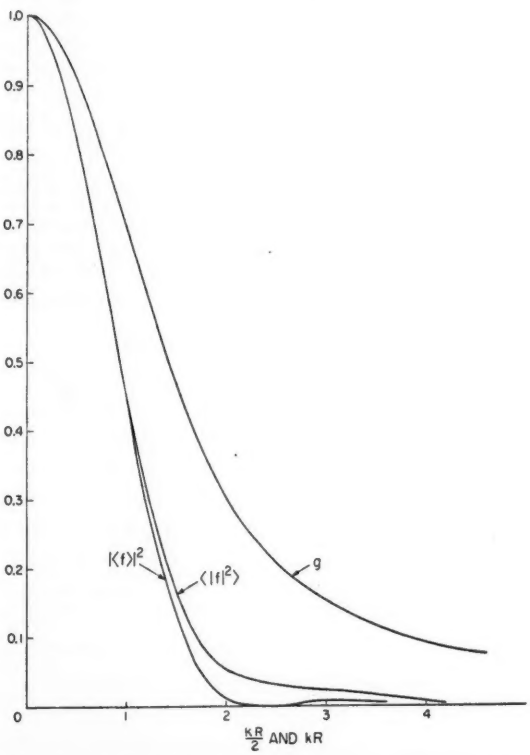


FIG. 4. The spherical averages $\langle|f|^2\rangle$ and $|\langle f\rangle|^2$ of f *vs* $\kappa R/2$ for Meckler's ground electronic wave function. Also plotted is the form g of the total magnetic cross section in center-of-mass coordinates *vs* kR.

TABLE III. The extreme values of A_l and the corresponding arguments for $l=0, 2, 4, 6, 8$ derived (a) from Meckler's Gaussian, and (b) from hydrogenic (see reference 15) $(Z=4.44)2p_{\pm}$ atomic orbitals.

	l	0	2	4	6	8
(a)	$(A_l)_{\text{ext}}$	1.00	0.09	−0.017	−0.009	−0.002
	$(\kappa R/2)_{\text{ext}}$	0.0	2	3.8	4.4	5
(b)	$(A_l)_{\text{ext}}$	1.00	0.11	−0.005	−0.0072	−0.003
	$(\kappa R/2)_{\text{ext}}$	0.0	2.2	3.2	4.8	6.5

change of internuclear distance is only 0.03), and also the couplings of the electronic spin leading to fine structure[17] in the rotational levels. One would hardly expect that if these were taken into account the change in scattering, even in center-of-mass coordinates, would be greater than that obtained if reasonable changes are made in the $2p_{\pm}$ orbitals.

V. TOTAL MAGNETIC CROSS SECTION

From the partial cross sections $\sigma_{J'J}(\kappa)$ of (12) one could in principle compute the distribution in angle and energy of the scattered neutrons in the laboratory frame of reference.[18] Here we shall be concerned with the total cross section in the laboratory frame derived from the static approximation expression (14) together with Meckler's wave function (15). In accordance with the assumption of nearly elastic scattering involved in the static approximation, we set $\kappa=2k\sin(\theta/2)$, where θ is the scattering angle. By integrating (14) over all solid angle, we obtain for the total cross section in center-of-mass coordinates, in units of $\frac{2}{3}S(S+1)\times[(m/m_p)(g_n e^2/2mc^2)]^2=0.3650\times10^{-24}$ cm^2, the expression

$$\sigma(k)=\int\langle|f(\kappa)|^2\rangle\sin\theta d\theta d\phi=(2\pi/k^2)\int_0^{2k}\langle|f|^2\rangle\kappa d\kappa \qquad (19)$$

$$\rightarrow 4\pi \quad \text{as} \quad k\rightarrow 0.$$

The required integration was carried out numerically in terms of the more convenient integration variable $\kappa R/2$. The resulting function is plotted in Fig. 4. The transformation of the cross section in center-of-mass coordinates to laboratory coordinates is based on the principle that the number of neutrons scattered per unit time is independent of the Newtonian reference frame. Averaging[19] over a distribution of molecular velocities appropriate to a gas temperature of 20°C led to values of the total cross section given in Table IV. The agreement with experiment[20] is to within the experimental error.

TABLE IV. Total magnetic cross section σ per molecule of oxygen gas for neutrons of incident wavelength λ. In the limit of infinite wavelength the relative motion of the neutrons and molecules depends only on the molecular motion, and σ/λ approaches a limit independent of λ, which for $T=20$°C is 0.361 barn/angstrom.

λ (angstroms)	1	2	3	4	5	7.5	10	15	20	30
σ (barns)	0.130	0.490	0.981	1.516	2.015	3.054	3.96	5.76	7.52	11.09

The author is indebted to Dr. H. Palevsky and Dr. L. van Hove for several interesting discussions concerning this work. A large part of the computation was done by Miss Edith Moss. I am pleased to express my appreciation to others of the Solid-State and Molecular Theory Group at M. I. T. and especially to Professor J. C. Slater for the stimulating atmosphere I have enjoyed during the course of this work.

[17] The over-all fine structure splitting for each rotational level of interest is about 2 cm^{-1} corresponding to a neutron wavelength of 15A. These splittings will not affect the scattering, at least to the extent that the split levels are equally populated, and that scattering corresponding to transitions among them may be regarded as elastic.

[18] For a detailed discussion of the required transformations see reference 4.

[19] The same procedure is used in references 12 and 1. In the present case the integrals were carried out numerically.

[20] H. Palevsky and R. M. Eisberg (to be published).

Internal and External Bremsstrahlung Accompanying the Beta Rays of P^{32}

K. LIDÉN AND N. STARFELT
Radiation Physics Department, University of Lund, Lund, Sweden
(Received August 3, 1954)

The spectral distribution and the total energy of the internal bremsstrahlung of P^{32} were measured from 20 to 1000 kev with a NaI(Tl) scintillation spectrometer. Special care was taken to diminish and correct for bremsstrahlung from the apparatus and the surroundings. The experimental curve agrees with theory from 20 to 50 kev and in the region 50–120 kev shows a weak bump in a smoothed curve; from 120 to 500 kev the curve lies 15 percent, at 800 kev 45 percent, and at 1000 kev 70 percent above a theory based on allowed β transition; thus, above 120 kev, the experimental curve disagrees with the theory for both allowed and forbidden β transitions. The total energy per β decay in the internal bremsstrahlung was found to exceed the calculated value by 21 percent.

Using the same experimental arrangements the external bremsstrahlung emitted when the P^{32} β rays are completely stopped in carbon, aluminum, iron, tin, and lead was also measured. For light elements the observed spectrum was found to agree closely with the theory by Bethe and Heitler as improved by Elwert, but for heavy elements and high energies to exceed the calculated values considerably, e. g., for lead at 500 and 1000 kev by a factor of 1.4 and 2.0, respectively.

I. INTRODUCTION

A. Internal Bremsstrahlung (IB)

ALTHOUGH Aston[1] is widely credited with the discovery of this weak continuous nuclear γ radiation, the first to produce convincing experimental evidence of such radiation were Stahel and Kipfer,[2] v. Droste,[3] and Gray and Hinds.[4] It is true that Aston discussed the possibility of IB from RaE, but an analysis of his absorption measurements strongly suggested that the intensity of the external bremsstrahlung (EB) from the lead absorption foils and from the platinum wire of the source exceeded that of IB by a factor of 5 or more.

The first to measure the P^{32} IB was Sizoo *et al.*,[5] but the investigation by Wu[6] gave more reliable results. Wu studied EB+IB for various elements with an integration ion chamber. Extrapolation to $Z=0$ gave $IB/(IB+EB_{Al})=\frac{1}{4}$. When the experimental difficulties are considered, it appears that the agreement, claimed by Wu, between the calculated and the experimental total IB energy must be accepted with caution.

With the poor energy resolution of the ion chambers and G-M tubes used in these earlier investigations, it was not possible to study the spectral distribution with any degree of precision. Novey,[7] Madansky and Rasetti,[8] and Bolgiano *et al.*[9] were the first to use scintillation spectrometers in the investigations of IB from P^{32}, Y^{91}, and RaE for photon energies <250 kev. Renard[10] applied the proportional counter technique in the region of 3–30 kev. All of these authors found the shape of the spectrum to agree fairly well with theory.

The theory of IB put forward by Knipp and Uhlenbeck[11] and by Bloch[12] for allowed β transitions, was extended by Chang and Falkoff[13] and Madansky *et al.*[14] for other transitions.

Recently Michalowicz[15] and Goodrich and Payne[16] published IB measurements of P^{32} up to 400 and 900 kev, respectively. The shape of Goodrich and Payne's experimental spectrum agreed well with the theory for the allowed β transition, but not with the first forbidden scalar type given by Chang and Falkoff.[13]

In an earlier study of the scintillation spectrometer method used in the measurements of continuous γ and x-radiation, we stressed the importance of proper correction of the experimental results. The measurements made in the present investigation of the P^{32} IB were corrected in the way described in this earlier paper[17] (LS).

B. External Bremsstrahlung (EB)

This type of x-rays, emitted by β rays being stopped in matter, was first observed with certainty by Gray,[18] who used β rays from RaE. Chadwick[19] and Hess and Lawson[20] also studied EB from RaE. Later Gray[21] measured the fraction of the β energy emitted as EB when the β rays from RaE were completely stopped

[1] G. H. Aston, Proc. Cambridge Phil. Soc. **23**, 935 (1927).
[2] E. Stahel and P. Kipfer, Helv. Phys. Acta **9**, 492 (1936).
[3] G. v. Droste, Z. Physik **104**, 335 (1936).
[4] J. A. Gray and J. F. Hinds, Phys. Rev. **49**, 477 (1936).
[5] Sizoo, Eijkman, and Groen, Physica **6**, 1057 (1939).
[6] Chien-Shiung Wu, Phys. Rev. **59**, 481 (1941).
[7] T. B. Novey, Phys. Rev. **84**, 145 (1951); **89**, 672 (1953).
[8] L. Madansky and F. Rasetti, Phys. Rev. **83**, 187 (1951).
[9] Bolgiano, Madansky, and Rasetti, Phys. Rev. **89**, 679 (1953).
[10] G. A. Renard, J. phys. et radium **14**. 361 (1953).
[11] J. K. Knipp and G. E. Uhlenbeck, Physica **3**, 425 (1936).
[12] F. Bloch, Phys. Rev. **50**, 272 (1936).
[13] C. S. Wang Chang and D. L. Falkoff, Phys. Rev. **76**, 365 (1949).
[14] Madansky, Lipps, Bolgiano, and Berlin, Phys. Rev. **84**, 596 (1951).
[15] A. Michalowicz, J. phys. et radium **15**, 156 (1954).
[16] M. Goodrich and W. B. Payne, Phys. Rev. **94**, 405 (1954).
[17] K. Lidén and N. Starfelt, Arkiv Fysik **7**, 427 (1954). This paper will be referred to as LS.
[18] J. A. Gray, Proc. Roy. Soc. (London) **A85**, 131 (1911); and **A86**, 513 (1912).
[19] J. Chadwick, Phil. Mag. **24**, 594 (1912).
[20] V. F. Hess and R. W. Lawson, Wien. Ber. **125**, 661 (1916).
[21] J. A. Gray, Phys. Rev. **19**, 430 (1922); and **25**, 237 (1925).

in lead, iron, or paper. Stahel and Kipfer[2] observed that the IB of RaE was slightly more penetrating than the EB from Pb+RaE; they reported a total intensity ratio of $IB/EB_{Pb}=1/7$. Von Droste[3] who used a GM detector, was the first to estimate the spectral distribution of the RaE EB. Using high pressure ion chambers and UX β rays Staheland Mass a[22] studied the Z-dependence of the total EB energy emitted and found it to agree with Bethe and Heitler's[23] theory.

Wu[6] found, that the shape of the EB spectrum of P^{32} was independent of Z and that, roughly speaking, the total EB energy measured agreed with that calculated according to Bethe and Heitler.[23]

An experimental study of the P^{32} EB spectrum of Pb was made by Siegbahn and Slätis,[24] who used a β spectrograph and Pb converter. They also observed the superimposed characteristic x-rays from Pb. We studied these K x-rays quantitatively with a NaI scintillation spectrometer.[25] Similar measurements have since been reported by Michalowicz.[15]

Wyard[26] found reasonable agreement between approximate calculations from the Bethe-Heitler theory and absorption measurements of the EB spectrum of the P^{32} β rays stopped in brass. This spectrum was analyzed more closely by Goodrich *et al.*[27] who used a NaI(Tl) scintillation spectrometer. Applying Wyard's approximation they found, that from 50 to 1100 kev the experimental and calculated curves agreed within 10 percent.

The theory of the EB at nonrelativistic energies was developed by Sommerfeld[28] and Elwert.[29] The relativistic case was treated by Bethe and Heitler[23] by the Born approximation. Heitler[30] pointed out that this theory does not give reliable results for high-energy photons and heavy elements. Considering the lack of accurate measurements of true EB spectra, we made a detailed study of the EB spectra of the P^{32} β rays stopped in various elements. The experimental results were compared with the Bethe-Heitler theory in its original form and with the improvement suggested by Elwert as well as with Wyard's spectrum.

A preliminary report of part of the present investigation has been published.[31]

II. EXPERIMENTAL MEASUREMENTS

In an earlier set of measurements a single-channel pulse analyzer as described by Elmore and Sands[32] was used in conjunction with an EMI 5311 photomultiplier and a NaI(Tl) crystal 1.5 cm in diameter and 1.5 cm thick. Later, for the purpose of comparison and control, some of the measurements were repeated with a new differential precision pulse-height analyzer, of the type described by Francis and Bell,[33] a Du Mont 6292 photomultiplier and the same crystal as before (Fig. 1). The crystal was optically coupled to the photocathode by a thin film of Dow Corning DC200 silicone oil (10^5 centistokes) and then enclosed hermetically in a 0.15-mm Al reflector coated on the inside with a thin film of MgO. The electronic components were connected to a mains stabilizer, the temperature of the laboratory was thermostatically controlled and the apparatus started several hours before measuring a spectrum. Despite these precautions, frequent checking of the energy calibration and the channel width was necessary to secure measurements of satisfactory accuracy.

In the IB measurements the direct β rays were absorbed in a 1.02-g/cm^2 beryllium disk halfway between the source and the detector, which were 16 cm apart. The cylindric hole in the lead collimator (50 mm thick and covered by 1.1 g/cm^2 Perspex) was 20 mm in diameter. The EB produced in the Perspex absorber and in the air by the β rays was almost completely eliminated by the lead collimator and the steel cylinder (wall thickness 8 cm) enclosing the photomultiplier. An estimation showed that even at the highest IB photon energy measured (1.0 Mev) such EB disturbance was well below 5 percent.

In order to avoid the K x-rays from the lead collimator, the hole was lined with 0.7 mm Sn+0.2 mm Cu and the detector side with 1.5 mm Sn+0.5 mm Cu.

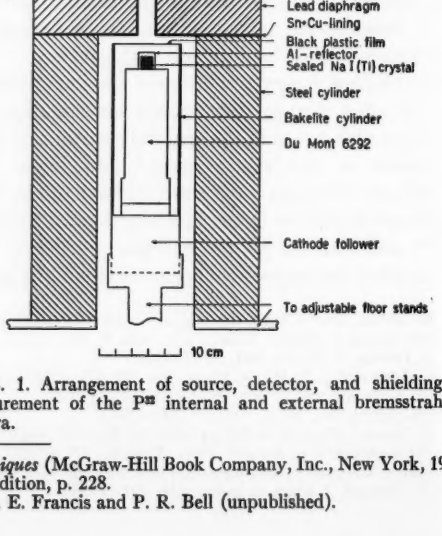


FIG. 1. Arrangement of source, detector, and shielding for measurement of the P^{32} internal and external bremsstrahlung spectra.

[22] E. Stahel and J. Massa, Helv. Phys. Acta **14**, 325 (1941).
[23] H. Bethe and W. Heitler, Proc. Roy. Soc. (London) **A146**, 83 (1934).
[24] K. Siegbahn and H. Slätis, Arkiv Mat. Astron. Fysik **34A**, No. 6 (1948).
[25] K. Lidén and N. Starfelt, Arkiv Fysik **7**, 193 (1954).
[26] S. J. Wyard, Proc. Phys. Soc. (London) **A65**, 377 (1952).
[27] Goodrich, Levinger, and Payne, Phys. Rev. **91**, 1225 (1953).
[28] A. Sommerfeld, Ann. Physik **11**, 257 (1931).
[29] G. Elwert, Ann. Physik **34**, 178 (1939).
[30] W. Heitler, *The Quantum Theory of Radiation* (Oxford University Press, London, 1954), third edition, Chap. V, Sec. 25.
[31] K. Lidén and N. Starfelt, Arkiv Fysik **7**, 83 (1954).
[32] W. C. Elmore and M. Sands, *Electronics Experimental Techniques* (McGraw-Hill Book Company, Inc., New York, 1949), first edition, p. 228.
[33] J. E. Francis and P. R. Bell (unpublished).

The P^{32} sources were prepared from absolutely calibrated P^{32} solutions, obtained from the Atomic Energy Research Establishment, Harwell, and containing about 5 mC/ml. The P^{32} was produced by $S^{32}(n,p)P^{32}$ and chemically processed. Before calibration, 10 μg/mC KH_2PO_4 was added as carrier to the originally carrier-free and chemically pure (no spectrographically detectable impurities) P^{32} solution (about 1 percent P^{33}). Using an infrared lamp a few drops were carefully evaporated on a 1.0-mg/cm^2 polystyrene film mounted on a Perspex ring of outer diameter 100 mm, inner diameter 90 mm, and thickness 0.5 mm, supported by three 2-mm Perspex rods. In this way a microscopically uniform spread of the sources (area 0.3–1 cm^2) was secured. Finally the sources were covered with 0.1 mg/cm^2 plastic films ("cocoon").

As the handling of a large source support on the balance offered some difficulties, the absolute amount of P^{32} was determined as follows. The bremsstrahlung emitted with the source between 1.0-g/cm^2 Al radiators was compared with the radiation from a similar source, prepared directly on an Al radiator by weighing and evaporating and covered with 1 mg/cm^2 polystyrene and another Al radiator.

The EB+IB spectra emitted from radiators of carbon, aluminum, iron, tin, and lead, which completely stopped the β rays from the P^{32} sources, were also measured.

The aforementioned IB geometry was also used for the EB measurements, which permitted most of the corrections to be made as for a parallel beam. All the IB and the EB spectra were studied with at least two different sources. The channel width employed varied between 4.5 kev for low photon energies and 80 kev for the highest energies measured. The intensity of the background spectrum, observed with the source removed, was low, 10 percent or less of the IB spectrum at 1000 kev. The statistical error for all experimental points was less than 3 percent. This required 8–10 hours' counting time per channel at the highest energies.

The energy calibration was frequently checked with a suitable set of the following lines: Na^{22} 1277 kev, Cs^{137} 662, Na^{22} 511, Au^{198} 412, I^{131} 364, Cr^{51} 330, Ce^{141} 145, Pb K_α 75 (P^{32}+Pb radiator), and Ce K_α 35 kev (P^{32}+Ce radiator).

III. CALCULATION OF THE TRUE SPECTRA FROM THE OBSERVED PULSE-HEIGHT DISTRIBUTIONS

A. Detector Corrections

The observed pulse-height distributions were corrected for background, dead-time of the analyzer, resolving power, Compton distribution, K x-ray escape, backscatter from the photomultiplier, and γ efficiency (LS).[17]

The ratio $k(E_\gamma)$ (LS, p. 450) between the number of pulses in the photopeak and the total number of observed pulses of a single γ-ray line with the energy E_γ was now determined with greater accuracy (Fig. 2). The result agrees within a few percent with values calculable from data given by Maeder *et al.*[34]

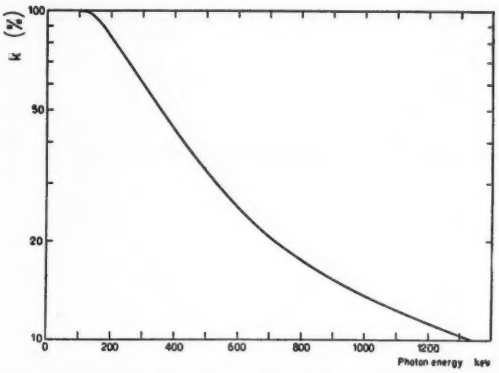


FIG. 2. The experimental ratio $k(E_\gamma)$ of the number of pulses in the photoelectric peak to the total number of pulses of a single γ ray line with the energy E_γ.

The correction for K x-ray escape and resolving power were made as follows. When corrected for K x-ray escape the distribution $N_3(\epsilon_\gamma)$ of the scintillations in the crystal gives the distributions $N_4(\epsilon_\gamma)$. A first approximation $N_4{}^I(\epsilon_\gamma)$ of $N_4(\epsilon_\gamma)$ was estimated. The corresponding distribution $N_3{}^I(\epsilon_\gamma)$ was then calculated. After that, using the experimentally determined resolving power, the resulting pulse-height distribution $N_2{}^I(\epsilon)$ was calculated according to LS. When $N_2{}^I(\epsilon)$ thus obtained deviates only slightly from the experimental curve, $N_2(\epsilon)$, the following equation is valid:

$$N_4(\epsilon_\gamma)=N_4{}^I(\epsilon_\gamma)N_2(\epsilon)/N_2{}^I(\epsilon).$$

Otherwise $HN_4{}^I(\epsilon_\gamma)$ was improved upon and the calculations were repeated. The experimentally determined line widths at half-height at 35 and 75 kev were 37 and 28 percent, respectively.

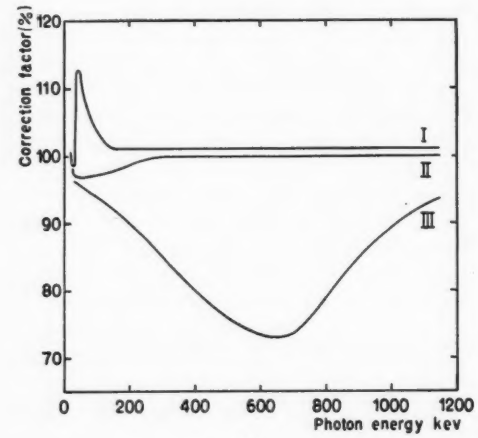
The calculated correction factors of the IB spectrum are shown in Fig. 3. The factors of the EB spectra were of roughly the same shape.

If two pulses ϵ' and ϵ'' coincide, the size of the resulting pulse ϵ is given by $\epsilon \leq \epsilon'+\epsilon''$. In a spectrum where the intensity decreases rapidly with increasing photon energy this effect can increase the counting rate at high energies considerably, if the total number of pulses is too high. In the present investigation with a pulse length of about 3 μsec this source of error was less than 0.5 percent.

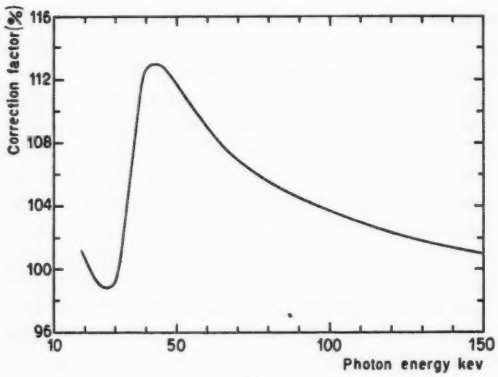
B. Correction for Scattering and Absorption

The contribution of scattered radiation from the Be absorber and from the hole of the collimator was calculated by the Klein-Nishina formula on the assumption that all scattering processes (single scat-

[34] Maeder, Müller, and Wintersteiger, Helv. Phys. Acta 27, 3 (1954).



(a)



(b)

FIG. 3(a). Correction factors of the scintillation pulse-height distribution of the internal bremsstrahlung of P^{32} for: (I) resolving power+K x-ray escape effect, (II) backscatter from the photomultiplier, (III) Compton electron distribution. (b). Correction factor (I) of Fig. 3(a) in the region 20–150 kev.

tering) in the hole occur only in a transverse plane of the wall exactly intermediate between the end openings (Fig. 4).

The IB and EB spectra were also corrected for absorption in the β ray absorber and the reflector (Fig. 4).

The true distribution of the P^{32} bremsstrahlung EB+IB, escaping perpendicularly from the radiator, is shown in Fig. 5. The lead spectrum was not corrected for resolving power below 150 kev.

C. Correction of the IB Measurements for EB

The IB pulse-height distribution includes EB emitted non-isotropically from the β absorber. Since the imperfection of the theory for the angular distribution of the thick target bremsstrahlung could not permit calculation of the corresponding correction, it was deduced from measurements obtained with β absorbers of Be, C, and Al. The values for $N_Z(E_\gamma)/\delta_Z$, where $N_Z(E_\gamma)$ is the measured photon intensity at energy E_γ and δ_Z the transmission of the absorber, were plotted as a function of Z for a number of energies. A linear extrapolation to $Z=0$ gave the correct value of the IB pulse-height distribution $N_{Z=0}(E_\gamma)$ for E_γ. The correction factor $N_{Z=0}(E_\gamma)/N_{Z=4}(E_\gamma)$ is given in Fig. 4; it decreased gradually from 7 percent at low energies to zero at 1000 kev.

The total correction factor for conversion of the pulse-height distribution into the true IB spectrum is given in Fig. 6. Above 200 kev its shape is dictated mainly by the effective photoelectric absorption of the crystal.

The P^{32} IB photon spectrum from 20 to 1000 kev is shown in Fig. 7.

D. Correction of the EB for Absorption and Scattering in the Radiator and for IB

It was assumed that all the photons emanate from the center of the source. The error introduced by this approximation was negligible at those energies where the absorption in the radiator was small and probably did not exceed 5 percent, even when the absorption was as much as 75 percent.

In the energy region around the peak at 75 kev in the bremsstrahlung spectrum of lead the correction for the absorption was more complicated, as this peak consists mainly of K x-rays, superimposed on a bremsstrahlung spectrum strongly attenuated by absorption in the radiator. These K x-rays are emitted not only after photoelectric absorption of bremsstrahlung photons but also after ionization in the K shell by β particles.[25] Because of the limited resolving power of the spectrometer this K x-ray peak could not be separated with sufficient accuracy from the continuous spectrum. This peak was instead regarded as a single line and was substituted by a Gaussian distribution which was then subtracted from the spectrum. In this way pulses caused by photons of

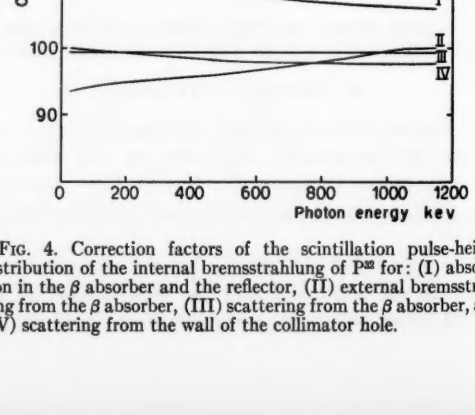


FIG. 4. Correction factors of the scintillation pulse-height distribution of the internal bremsstrahlung of P^{32} for: (I) absorption in the β absorber and the reflector, (II) external bremsstrahlung from the β absorber, (III) scattering from the β absorber, and (IV) scattering from the wall of the collimator hole.

energies below the K absorption edge of lead were removed from the spectrum above 100 kev.

The probability of a photon being scattered in the radiator towards the crystal was calculated for various primary photon directions according to the Klein-Nishina formula. The spectral distribution of the singly scattered radiation originating in the continuous spectrum under consideration was calculated by a procedure similar to that used in the calculation of the Compton electron distribution in the crystal.

The EB spectra (Fig. 9) were then obtained by subtraction of the true IB spectrum.

IV. SURVEY OF THE ERRORS

The accuracy of the measurements and of the calculations of the corrections for the IB are apparent from Table I. As to the spectra of EB+IB the various errors were approximately of the same order; in addition to the errors presented in Table I those in the calculations of scattering and absorption in the radiator must be taken into account. The calculated total rms deviation (Table I) does not include the statistical error and the error in the absolute calibration (2 percent).

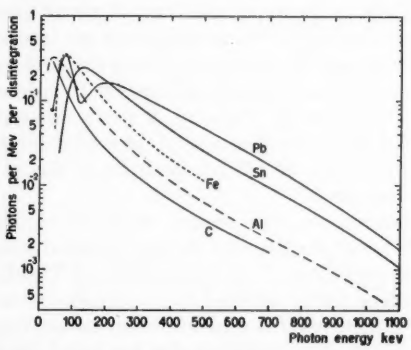


FIG. 5. The true spectra of the P^{32} bremsstrahlung photons escaping perpendicularly from radiators of about 1.0 g/cm² enclosing the source; number of photons per β disintegration per 1-Mev energy internal *vs* photon energy.

The error of the calculation of the Compton electron continuum was due mainly to the approximation of the Compton electron distribution for a certain γ-ray energy E_γ to a straight line. The estimation of the error was based on calculations performed using a slope of this line representing the greatest possible deviation from the approximation used.

The error in the crystal efficiency was due mainly to the inaccuracy of the measurement of $k(E_\gamma)$ caused above all by the radiation scattered from the collimator and the photomultiplier. The values, according to White,[35] of the absorption coefficients of sodium and iodine were taken as correct, but, as pointed out by Maeder *et al.*,[34] the possibility of systematic errors cannot be excluded.

[35] G. R. White, Natl. Bur. Standards Rept. No. 1003, May 13, 1952 (unpublished).

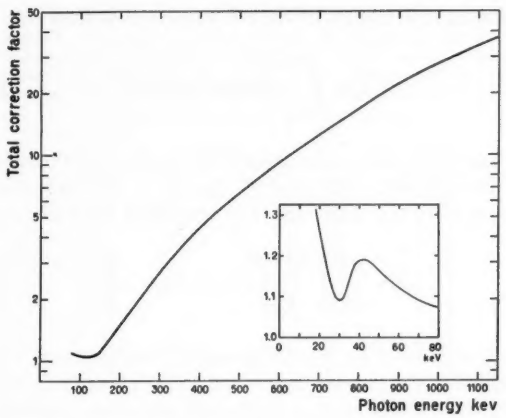


FIG. 6. The total correction factor converting the scintillation pulse-height distribution of the internal bremsstrahlung of P^{32} into the true spectrum.

Some of the aforementioned errors (backscatter, crystal efficiency, EB disturbance) can be diminished by using better geometry. However, the geometry is determined by the strength of the source which cannot be increased too much because of EB then produced in the source.

The resolving time of the pulse analyzer and the related maximum pulse rate is also a limiting factor. Finally, it may be mentioned that the accuracy of the measurements may be further improved upon by the use of such electronic components in the spectrometer as provide perfect long-term stability.

V. RESULTS AND DISCUSSION

A. Internal Bremsstrahlung

According to Knipp and Uhlenbeck's theory,[11] β decay and radiation are regarded as independent processes. The probability $S(k)dk$ for the emission of a

TABLE I. Sources of error in the P^{32} internal bremsstrahlung spectrum (in percent). Errors in the corrections for backscatter (<1 percent), external bremsstrahlung from the β absorber (1 percent), and scattering in the collimator (<1 percent) are included in the total error, which does not, however, include counting statistics or the inaccuracy of the absolute calibration of the source (2 percent).

E_γ kev	Channel width	Energy calibration	Resolving power + K x-ray escape	Compton electron continuum	Crystal efficiency	Total (rms)	Counting statistics
30	2	1	2	0.5	0.5	3.3	1.5
50	2	1	1	0.5	0.5	2.9	2
100	2	0.5	0.5	0.5	1	2.6	2
200	2	0.5	0.5	0.5	2	3.2	3
500	1	0.5	0.5	2	3	4	3
700	1	1	0.5	5	3	6	3
1000	1	3	0.5	3	4	6	3

photon with energy between k and $k+dk$ is then

$$S(k)=\int_{1+k}^{E_m} P(E_0)\Phi(E_0,k)dE_0, \qquad (1)$$

where $P(E_0)dE_0$ is the probability for the emission of a β particle with total energy E_0; E_m is the end-point energy of the β spectrum. The expression for Φ is

$$\Phi=\frac{\alpha p}{\pi p_0 k}\left[\frac{E_0^2+E^2}{E_0 p}\ln(E+p)-2\right], \qquad (2)$$

where $E=E_0-k$; p_0 and p are the momenta of the radiating electron corresponding to the energies E_0 and E, respectively, and α is the fine structure constant, 1/137. All energies are in units of mc^2 and all momenta in units of mc. However, Chang and Falkoff[13] pointed out "that there appears to be no good reason to suppose that the probability for the overall process should be simply the product of the probability for each," but they prove that it yields exactly the same result as the more rigorous second-order perturbation method for allowed transitions. In both methods Born's approximation is used, and the effect of the nuclear charge on the radiation is thus neglected. According to Novey,[7] Hellund has shown that this effect is unimportant in the low-energy photon region investigated by Novey. However, if the experimentally determined β spectrum is used as $P(E_0)$ in (1), the effect of the nuclear charge is, to a certain extent, taken into consideration.

The solid curve in Fig. 7 shows $S(k)$ calculated from (1) by taking $P(E_0)$ from the β spectrum of P^{32} according to Jensen *et al.*[36] The experimental points are the result of four runs with the same geometry, using two P^{32} sources of 2.5 and 1.0 mC, prepared from two different Harwell solutions.

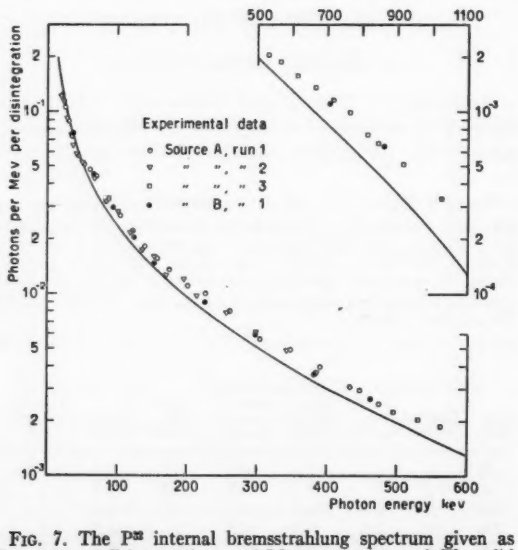


FIG. 7. The P^{32} internal bremsstrahlung spectrum given as photons per β disintegration per 1-Mev energy interval. The solid line represents the theoretical spectrum according to Knipp and Uhlenbeck (see reference 11). The statistical error is smaller than the size of the points. Source $A=2.5$ mC; source $B=1.0$ mC.

From 20 to 50 kev the experimental and calculated curves agree with an accuracy of ± 5 percent. The weak bump in the region 50–120 kev with its maximum about 15 percent above a smoothed curve is more readily recognized in an IB energy *vs* photon energy representation. This bump is much weaker than in Novey's IB spectrum,[7] because he did not eliminate the Pb K x-rays from the collimator hole. Such a bump can also be discerned in the spectra given by Bolgiano *et al.*[9] and by Michalowicz.[15] It is as yet not possible to say whether the bump is ascribable to an impurity or to P^{32} itself. Apart from this, no evidence of nuclear γ rays was found. From 120 to 500 kev the shapes of the experimental and calculated spectra show close agreement but disagree definitely above 600 kev. On an absolute basis the experimental points in the region 120–500 kev lie about 15 percent above theory, at 800 kev 45 percent, and at 1000 kev 70 percent. These differences are inconsistent with the experimental accuracy. Since the same general results have been obtained here[37] for the IB spectrum of S^{35} (end point energy 168 kev), it seems reasonable to suppose, that the improvement of the theory worked out by Nilsson[38] for the S^{35} IB spectrum is also of some importance for P^{32}.

In contrast to the present result Goodrich and Payne[16] found the shape of the experimental spectrum to agree with theory from 100 to 900 kev. The discrepancy can be partly ascribed to an additional contribution of low-energy EB photons, from the air, from the Lucite β absorber, and from the 5-mg/cm^2 P^{32} source used by these authors. Furthermore, the experimentally measured spectral distribution of the EB contribution from the β absorber differs from ours.

In this connection it might be mentioned that earlier measurements of the shape of the P^{32} IB spectrum with poorer geometry described elsewhere[25] showed good agreement with theory from 50 to 800 kev, but the experimental spectrum was about 40 percent above the calculated. Later, closer analysis indicated, that there was a greater amount of EB from the Perspex β absorber than that found before. As the ratio EB/IB is greatest at low energies, this additional contribution happened to smooth out the shape of the IB spectrum with consequent good agreement with theory.

In other papers of the P^{32} IB spectrum, only the low-energy end was studied. The results reported by Novey[7] and by Bolgiano *et al.*[9] were discussed to some extent in LS. Furthermore, these authors as well as Michalowicz[15] underestimated the EB contribution from the

[36] Jensen, Nichols, Clement, and Pohm, Phys. Rev. **85**, 112 (1952).
[37] N. Starfelt and N. L. Svantesson, Phys. Rev. **97** (1955) (in press).
[38] S. B. Nilsson (to be published).

β absorber in assuming isotropic distribution (Sauter,[39] Sesemann[40]). Novey and Michalowicz ignored the Compton electron distribution, the back-scatter from the photomultiplier, and the energy dependence of the geometrical crystal efficiency, which may easily give errors of anything up to 15 percent.

It may also be mentioned, that the good agreement from 3 to 30 kev reported by Renard[10] is based upon calculated values, which are about 50 percent smaller than those calculated by Novey, Michalowicz, and by us.

The present measurements gave a total amount of IB energy per β disintegration of 2.87×10^{-3} mc^2/β, i.e., 21 percent greater than the calculated value, 2.37×10^{-3} mc^2/β. The result reported by Bolgiano *et al.*[9] agrees with theory, but their value was estimated only from measurements of all pulses above 90 kev. Michalowicz[15] obtained a value 10 percent above theory in the region 0–400 kev, in agreement with our measurements.

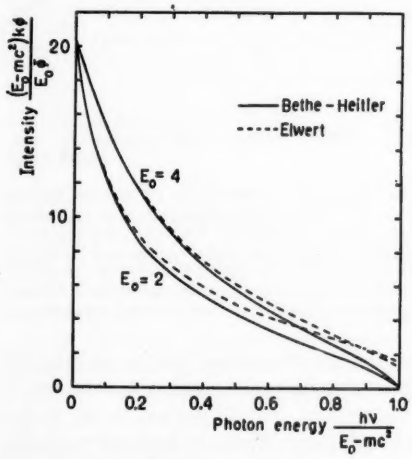


FIG. 8. Energy distributions $k\phi(E_0-mc^2)/E_0\bar{\phi}$ of the thin-target bremsstrahlung of monochromatic electrons with total energy $E_0=2mc^2$ and $E_0=4mc^2$ as a function of $k/(E_0-mc^2)$, calculated for lead according to Bethe and Heitler (see reference 23) (solid lines) and according to Elwert (see reference 29) (dotted lines). Here k is the photon energy in units of mc^2, ϕ the cross section integrated over all photon directions for the emission of a photon with energy between k and $k+dk$ $\phi=Z^2r_0^2/137$, and r_0 the classical electron radius.

B. External Bremsstrahlung

The cross section ϕ of a single atom (thin target) for the emission of an EB photon with energy between k and $k+dk$, integrated over all photon directions, was calculated by Bethe and Heitler[23] using Born's approximation, which is valid for low Z and relativistic velocities of the radiating electron before and after the process. For nonrelativistic energies Elwert[29] and later Guth[41] gave a correction factor, $f(\xi,\xi_0)$, to ϕ

$$f(\xi,\xi_0)=\frac{\xi}{\xi_0}\frac{1-e^{-2\pi\xi_0}}{1-e^{-2\pi\xi}},\quad \xi=\frac{Ze^2}{hv},\quad \xi_0=\frac{Ze^2}{hv_0},\tag{3}$$

where v and v_0 are the velocities of the electron before and after the process. Furthermore, they proposed as a plausible assumption this factor to hold also for relativistic energies.

The cross section ϕ for lead calculated both according to Bethe-Heitler and to Elwert is shown in Fig. 8 with the intensity $k\phi(E_0-mc^2)/E_0\bar{\phi}$ as a function of the photon energy $h\nu/(E_0-mc^2)$.

According to Bethe and Heitler[23] the EB spectrum, $n(k,E')$, obtained when completely stopping an electron of total primary energy E' is

$$n(k,E')=\int_{1+k}^{E'}\frac{N\phi\,dE_0}{(-dE_0/dx)},\tag{4}$$

where N is the number of atoms per cm^3 and $-dE_0/dx$ the energy loss of the electron per cm. A set of values of $n(k,E')$ for various k was calculated by numerical integration of (4), using values of $-dE_0/dx$ calculated according to Aron *et al.*[42] with values of the mean excitation potential from Segrè.[43] The P^{32} EB spectrum $S(k)$ at complete absorption of the β rays is then

$$S(k)=\int_{1+k}^{E_m}P(E')n(k,E')dE'.\tag{5}$$

Numerical integration of (5) for various k gave $S(k)$ for the elements investigated (Table II).

Here the contribution of the atomic electrons to the EB was not taken into account. A proper consideration of this effect would probably increase the calculated

TABLE II. The P^{32} external bremsstrahlung photon spectra for complete stopping of the beta rays in C, Al, Sn, and Pb, calculated according to the theories of Bethe and Heitler and of Elwert, respectively (references 23 and 29). After multiplication by 10^{-3} the figures give the number of photons per Mev per disintegration.

Energy Mev	Carbon Elwert	Aluminum Elwert	Aluminum Bethe-Heitler	Tin Elwert	Lead Elwert	Lead Bethe-Heitler
0.0511	160	380		1760	3000	
0.1022	54.7	132	114	603	1016	844
0.1533	27.4	65.1	56.5	298	509	430
0.2555	9.54	23.3	20.9	107.2	184	155
0.358	4.33	10.5	9.55	49.4	85.5	68.6
0.511	1.50	3.80	3.45	17.8	30.6	24.3
0.639	0.697	1.73		8.28	14.0	
0.767	0.321	0.810	0.670	3.93	6.87	4.75
0.894	0.140	0.350		1.76	3.07	
1.022	0.057	0.144	0.120	0.76	1.33	0.82

[39] F. Sauter, Ann. Physik **20**, 404 (1934).
[40] G. Sesemann, Ann. Physik **40**, 66 (1941).
[41] E. Guth, Phys. Rev. **59**, 325 (1941).
[42] Aron, Hoffman, and Williams, U. S. Atomic Energy Commission Report AECU-663, May 28, 1951 (unpublished).
[43] E. Segrè, *Experimental Nuclear Physics* (John Wiley & Sons, Inc., New York, 1953), Vol. 1, p. 203.

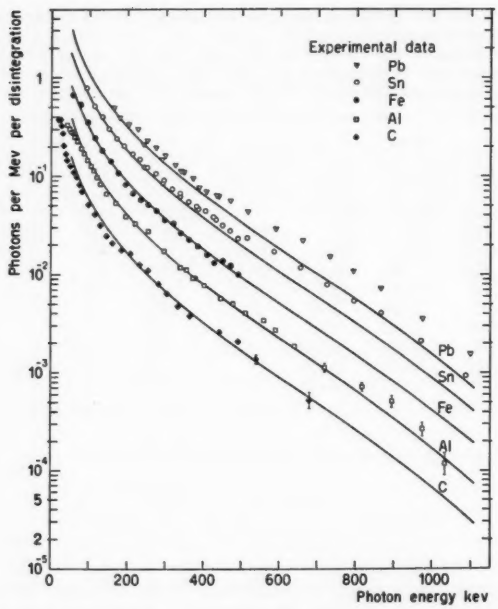


FIG. 9. The P^{32} external bremsstrahlung spectra produced when completely stopping the β rays in radiators of C, Al, Fe, Sn, and Pb, given as photons per β disintegration per 1-Mev energy interval. The solid lines are calculated according to Elwert (see reference 29).

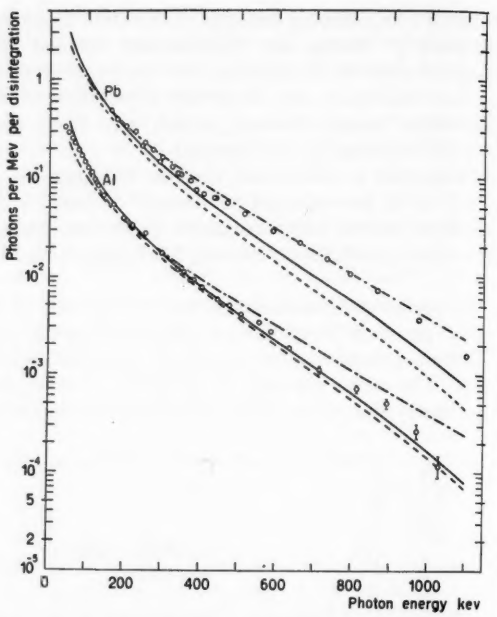


FIG. 10. The photon spectra of the P^{32} external bremsstrahlung from Al and Pb. The circles represent the experimental points, the solid lines and the dotted lines the calculated spectra according to Elwert (see reference 29) and to Bethe and Heitler, (see reference 23) respectively. The dot dashed lines represent the result of Wyard's (see reference 26) approximation of the Bethe and Heitler theory; it is normalized to the experimental spectrum at 0.5 mc^2.

bremsstrahlung intensity by a factor of the order of $1+1/Z$, at least for extreme relativistic energies.[30]

The solid curves in Fig. 9 represent $S(k)$ according to Elwert for C, Al, Fe, Sn, and Pb and the various sets of points the respective experimental spectra.

The EB spectra of the elements C, Al, and Fe agree almost exactly with theory in the photon energy region studied. Generally, smoothed curves through the experimental points deviate less than 10 percent. Being calculated as differences between two large numbers, EB+IB and IB, the aluminum points above 800 kev are less accurate.

As to Sn and Pb, the disagreement between experiment and theory above 300 kev increases with energy. At 300 kev the difference is 8 percent for Sn and 15 percent for Pb, at 600 kev 43 and 58 percent, and at 1000 kev 90 and 95 percent, respectively.

For comparison, the solid curves in Fig. 10 show $S(k)$ of Al and Pb according to Elwert, the dashed curves $S(k)$ according to Bethe and Heitler, and the circles the experimental points. Concerning Al, the two theories differ but slightly, only 15 percent in the whole energy region investigated, while the lead curves diverge with increasing energy. Thus the Elwert-factor displaces Bethe and Heitler's EB spectrum in the desired direction, but the discrepancy is still considerable. For electron energies between nonrelativistic and extreme relativistic values the EB cross section for high photon energies and large Z comes out too small in the theories discussed above.

The dot-dashed curves in Fig. 10, normalized to the experimental points at 0.5 mc^2, were calculated from the rather rough approximation of the Bethe-Heitler-Sommerfeld EB spectrum recently proposed by Wyard.[26] This approximation gives the shape of the EB spectrum of Pb surprisingly well up to 900 kev, but deviates at higher energies (the EB of Pb was also measured at 1217 and 1330 kev, not included in Figs. 9 and 10). However, the representation of the EB spectrum of Al by Wyard's spectrum is less successful.

From Fig. 9 it is obvious that the practically linear relation between the EB intensity and the atomic number Z, which follows from the Bethe-Heitler theory, is only approximately fulfilled; the deviation increases with increasing photon energy and increasing Z.

The observed total EB energy, EB_{obs}, as well as the ratio EB_{obs}/EB_{calc}, where EB_{calc} is calculated according

TABLE III. The experimentally determined energy, EB_{obs}, of the P^{32} external bremsstrahlung of various elements completely stopping the β rays, and the ratio EB_{obs}/EB_{calc}, where EB_{calc} is the calculated external bremsstrahlung according to Elwert (see reference 29).

Element	C	Al	Fe	Sn	Pb
Energy region, Mev	0−0.5	0−1.7	0−0.5	0−1.7	0.1−1.7
EB_{obs}, mc^2/β	0.00310	0.00888	0.0173	0.0471	0.0558
EB_{obs}/EB_{calc}	0.896	0.946	0.944	1.119	1.301

to Elwert, is given in Table III. As appears from Table III, the deviation from theory increases with increasing atomic number of the radiator element.

In the literature there is a lack of accurate observations of EB spectra excited by β rays. Goodrich *et al.*[27] investigated the shape of the EB+IB spectrum from P^{32} surrounded by 1.1-g/cm^2 brass, and found agreement within 10 percent from 50 to 1100 kev with Wyard's approximation discussed above. Michalowicz[15] has recently investigated the P^{32} EB spectrum of various elements in connection with corrections to a measurement of the IB spectrum. He made no comparison with theory, and details of the experimental results are given only for the K x-ray peaks.

The present investigation was carried out mainly at the Department of Physics, University of Lund. The authors wish to express their gratitude to Professor B. Edlén and Professor S. v. Friesen for their interest and support in this work and to J. Cederlund, B. Forkman, and N. L. Svantesson for valuable assistance.

PHYSICAL REVIEW VOLUME 97, NUMBER 2 JANUARY 15, 1955

Excitation Function for Na^{22} from Deuterons on Aluminum*

S. O. Ring and L. M. Litz
Livermore Research Laboratory, California Research and Development Company, Livermore, California
(Received May 3, 1954)

The cross section for the reaction $Al^{27}(d,d\alpha n)Na^{22}$ has been determined as a function of deuteron energy in the energy range from 30 to 190 Mev. This reaction may be used for monitoring a high-energy deuteron beam.

INTRODUCTION

In order to calculate cross sections of cyclotron-induced activities, it is necessary to know the integrated beam intensity, i.e., the number of particles impinging on the target. This intensity can be obtained by the use of a monitor material which undergoes a nuclear reaction for which the cross section as a function of energy is accurately known.

A standard monitor reaction used in deuteron bombardment work is $Al^{27}(d,\alpha p)Na^{24}$. The disintegration rate of radioactive Na^{24} obtained by counting the thin aluminum foils is used to monitor the beam. Na^{24} is a 15-hour, β^- emitter.[1] In those cases where the monitor foils cannot be counted soon after bombardment a monitor with a longer half-life would be desirable. The reaction $Al^{27}(d,d\alpha n)Na^{22}$ occurs simultaneously with the reaction given above. Na^{22} is a 2.6-year[1] positron emitter.

In the following experiments, thin aluminum foils were irradiated with deuterons in the 184-inch Berkeley cyclotron. The aluminum foils were counted and the cross section for Na^{22} was obtained as a function of deuteron energy by using the Na^{24} as an internal monitor.

EXPERIMENTAL

For the irradiations, aluminum foils 0.002-inch thick were used. A one-half mil aluminum foil was placed on either side of the 0.002-inch foil to eliminate any errors caused by recoils. These aluminum stacks were placed in a target holder mounted on the end of the probe of the 184-inch cyclotron. Various deuteron energies were obtained by varying the distance of the foils from the center of the cyclotron tank. The foils were mounted in such a way that the deuteron beam hit perpendicular to the surface of the foils so that a single energy was obtained at each irradiation. (The attention in energy in passing through such a thin target is negligible.) Twelve irradiations were made with two separate aluminum target foils at each irradiation energy. The irradiations were from 15 to 30 minutes with a deuteron beam of approximately 0.5 microampere. After irradiation, each foil was cut in half and both pieces mounted side by side on a thick aluminum backing plate so as to give a source of small area. These were counted under constant geometry with an Amperex type 100C, end-window Geiger tube. The tube and sample were contained in an aluminum-lined, lead counting chamber.

RESULTS AND DISCUSSION

The cross section as a function of energy for Na^{24} from aluminum bombarded by deuterons has been accurately established.[2] The activity of the aluminum foils was followed until the activity from the Na^{24} had died out leaving the activity caused by Na^{22} alone. (No other activities are formed with half-lives greater than several minutes.) The activities at the end of the bombardment were obtained for the two isotopes by resolution and extrapolation of the decay curves. Since the foils were counted under constant geometry

* This work was carried out under a U. S. Atomic Energy Commission contract.

[1] Hollander, Peckman, and Seaborg, Revs. Modern Phys. 25, 469 (1953).

[2] Batzel, Crane, and O'Kelley, Phys. Rev. 91, 939 (1953).

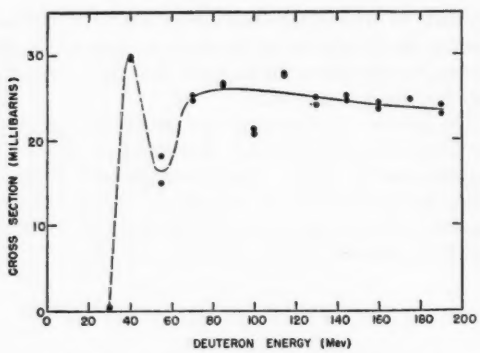


FIG. 1. Cross section of Na^{22} from aluminum as a function of deuteron energy.

conditions, it was only necessary to know the relative counting efficiency of Na^{22} and Na^{24} in order to calculate the Na^{22} cross sections. Na^{22} emits 0.54-Mev positrons and Na^{24} decays with the emission of a 1.4-Mev negatron.[1] The contribution to the activity from the gamma rays emitted by each isotope was small, less than two percent, when counted on a chlorine-argon beta counter and was neglected for monitoring purposes. The total air and window absorbing thickness was 5 milligrams per square centimeter. The correction factor for the air and window absorption for Na^{22} was taken as 1.08 and for Na^{24} as 1.04. These values were determined from absorption curves obtained for each activity. The correction for self-absorption and self-scattering were taken from Nervick's work[3] and were 1.05 for Na^{22} and 1.10 for Na^{24}. Backscattering corrections were taken as 1.28 for Na^{24} from the work of Burtt,[4] and 1.22 for Na^{22} from unreported work done in this laboratory.

The values of the Na^{22} cross section are plotted in Fig. 1 as a function of deuteron energy. In order to obtain low energies in the 184-inch cyclotron, the targets had to be placed close to the center of the tank, where there is an appreciable flux of neutrons with sufficient energy to cause the reaction $Al^{27}(n,\alpha)Na^{24}$. As a result of this reaction, cross sections for Na^{24} below 50 Mev are higher than they should be for deuterons alone.

The cross sections of Na^{22} in this low-energy range are therefore too low. That part of the curve represented by the broken line is only a guess.

The authors would like to express their appreciation to the crew of the 184-inch cyclotron at Berkeley who performed the necessary irradiations.

[3] W. E. Nervik and P. C. Stevenson, University of California Radiation Laboratory, Unclassified Report UCRL-1575, 1951 (unpublished).
[4] B. P. Burtt, Nucleonics 5, No. 2, 28 (1949).

PHYSICAL REVIEW VOLUME 97, NUMBER 2 JANUARY 15, 1955

Energy Levels in Ca^{42} and Ca^{44}*†

JOHN P. SCHIFFER‡
Sloane Physics Laboratory, Yale University, New Haven, Connecticut
(Received August 4, 1954)

The $K^{39}(\alpha,p)Ca^{42}$, $K^{41}(\alpha,p)Ca^{44}$, and $Ca^{42}(d,p)Ca^{44}$ reactions were studied and ground state Q-values of −0.19, 0.98, and 9.07 Mev respectively were found. Energy levels in Ca^{42} were found at 1.51, 1.95, 2.29, 2.59, and 3.02 Mev and in Ca^{44} at 1.13, 1.92, 2.28, 2.58, 2.97, and 3.17 Mev. An anomalously low cross section was observed for the $Ca^{42}(d,p)Ca^{44}$ reaction.

I. INTRODUCTION

PRELIMINARY work on the $K^{39}(\alpha,p)Ca^{42}$ reaction has been reported previously.[1] Since that time an increase in the cyclotron beam currents available and the installation of a magnetic beam analyzer have made a more accurate study of this reaction possible. Work was also done on reactions providing information on the level structure of Ca^{44}.

* Part of a dissertation presented to the Graduate School of Yale University in partial fulfillment of requirements for the degree of Doctor of Philosophy.
† Assisted by the joint program of the Office of Naval Research and the U. S. Atomic Energy Commission.
‡ Present address: Physics Department, Rice Institute, Houston, Texas.
[1] J. P. Schiffer and E. Pollard, Phys. Rev. 91, 474(A) (1953).

II. EXPERIMENTAL METHOD

The 4.15-Mev deuteron beam and the 8.22-Mev alpha-particle beam of the Yale cyclotron were used in these investigations. Thin targets of the isotopes under study were placed in the analyzed particle beam. Protons from the target were slowed down in aluminum absorbers and then detected in argon filled proportional counters. "Peaking" of the group structure was achieved by setting the discriminator circuit biases to permit only the highest pulses to be recorded. These were caused by particles stopping at the end of the counter and thus causing maximum ionization at the end of their range. The recorded pulses thus provided a differential absorption curve of the proton spectrum. The

electronic circuits used in this method have been described previously by several investigators from this laboratory.[2,3]

Targets were evaporated on Au or Ta backing. Instead of using pure K or Ca targets, it was found more convenient to use KI and CaI, these compounds being more stable under atmospheric conditions than the pure metals.

III. $K^{39}(\alpha,p)Ca^{42}$

Early work on this reaction was done by Pollard and Brasefield[4] with natural alpha sources. With very low yields the Q-values obtained provided more of a qualitative indication than an exact assignment. Siegbahn observed a gamma ray at 1.51 Mev following the beta decay of K^{42}, which was assumed to correspond to the first excited state of Ca^{42}. Stevenson and Deutsch measured the beta-gamma angular correlation for this transition and concluded that the spin and parity of the level was 2^+ in agreement with shell theory predictions for the first excited states of even-even nuclei.

Work on the (α,p) reaction indicated that the state in question was indeed the first excited state of Ca^{42}. The ground state Q-value was found to be -0.19 ± 0.07 Mev. The uncertainty was due to the inaccuracy with which the range of a proton group could be established. Nonuniformity in commercial foils has been found to be a large potential source of error in absorption measurements by some investigators.[5] No large variations were found in the commercial aluminum foils used, but a slight nonuniformity was detected in some gold foils. A further source of error was the lack of accuracy with which the stopping power of gold was known in the energy range studied. These considerations apply to the assignment of errors to all subsequent data.

A typical curve for this reaction is shown in Fig. 1. In calculating the levels of excitation in the nucleus from a curve such as this the errors tend to be smaller than those for the Q-values because when differences between Q-values are taken some of the systematic errors cancel. The values obtained for excited states in

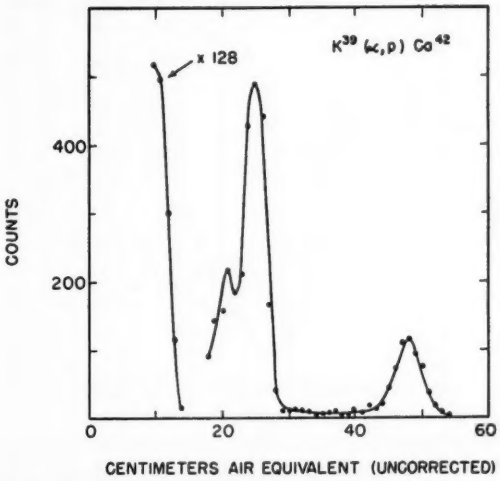


FIG. 1. The proton yield at 0° from $K^{39}(\alpha,p)Ca^{42}$. The peak on the left is that due to elastically scattered protons. Other runs, not shown here gave group structure at higher excitation.

Ca^{42} are shown in Table I. Level assignments above 3 Mev, while undoubtedly corresponding to levels in the nucleus, do not necessarily represent all the levels present. Runs at two different beam energies showed an entirely different shape for the proton spectrum in this region. Levels at lower excitation also changed in their relative intensities but no indication of new levels of energy lower than 3.3 Mev was found. It can only be concluded that the level density at higher excitation is greater than the resolution of the equipment used. No particular attempts were made in this investigation to measure levels higher than about the fourth excited state.

The Q-values cited are the result of six independent runs. While some reduction of experimental error would be possible with the accumulation of more data, the present method does not permit much greater accuracy.

TABLE I. Energy levels from the $K^{39}(\alpha,p)Ca^{42}$ and $K^{41}(\alpha,p)Ca^{44}$ reactions.

Ca^{42}	Ca^{44}
1.51±0.05	1.13±0.05[b]
1.95±0.07	1.92±0.05
2.29±0.05	2.28±0.05[b]
2.59±0.07	2.58±0.05
3.02±0.05	2.97±0.05
3.30±0.10[a]	3.17±0.05
3.75±0.07	
4.09±0.10[a]	

[a] Observed only with a beam energy of 7.75 Mev.
[b] Observed also in the $Ca^{43}(d,p)Ca^{44}$ reaction.

IV. $Ca^{43}(d,p)Ca^{44}$

Information on the levels structure of Ca^{44} has been known for some time. Hibdon, Kurbatov, and Pool[6] measured the gamma-ray energy following beta decay of Sc^{44} and found it to be 1.33 Mev. Later work found this energy to be 1.18[7] and 1.16[8] Mev. Recent work by Cohen[9] on the decay of K^{44} found evidence of 1.13, 2.07, 2.48, and 3.60 Mev gamma rays. It was also suggested from the shape of the beta decay curve of Sc^{44} that the ground state of Ca^{44} was even in parity and that the 1.13-Mev state was also even, with the gamma ray going by electric quadrupole or magnetic dipole transition. This is consistent with the shell theory prediction

[2] W. O. Bateson, Phys. Rev. **80**, 982 (1952).
[3] A. B. Martin, Phys. Rev. **71**, 127 (1947).
[4] E. Pollard and C. J. Brasefield, Phys. Rev. **50**, 890 (1936).
[5] S. K. Allison and S. D. Warshaw, Revs. Modern Phys. **25**, 779 (1953).
[6] Hibdon, Pool, and Kurbatov, Phys. Rev. **67**, 289 (1945).
[7] W. H. Cuffey, Phys. Rev. **79**, 180 (1950).
[8] J. Bruner and L. M. Langer, Phys. Rev. **79**, 606 (1950).
[9] B. L. Cohen, Phys. Rev. **94**, 117 (1954).

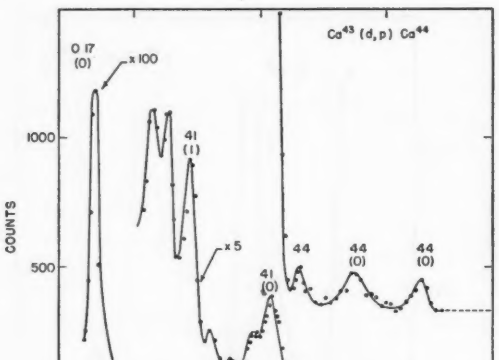


FIG. 2. The proton yield at 90° from $Ca^{43}(d,p)Ca^{44}$. The target contained 23.4 percent Ca^{40}, 1.6 percent Ca^{42}, 68 percent Ca^{43} and 7 percent Ca^{44}. The numbers above the peaks refer to the product nucleus. The peaks between the ground state and first excited state of Ca^{41} are due to Ca^{43} and Ca^{45}. The oxygen peak is shown for calibration purposes. The dotted line represents the neutron background.

for even-even nuclei to have 0^+ ground states and 2^+ first excited states.

The measurement of levels in Ca^{44} was undertaken by two means. It was thought that the (d,p) reaction on Ca^{43} would be comparable in yield to the $Ca^{40}(d,p)Ca^{41}$ reaction and therefore it was tried first. The Q-value calculated from mass values was 9.20 Mev which was very favorable. Initial attempts, however, failed to show any group structure high enough in energy to be associated with this Q-value.

Because of the large background in counting rates caused by the fast-neutron background associated with the deuteron beam of the cyclotron, a proportional counter telescope was first tried. Pulses from the first counter were used without bias to trigger one-half of a coincidence circuit, the other half of which was triggered by "peaked" counts in the second counter. A target containing 68 percent Ca^{43} and 23 percent Ca^{40} [10] showed group structure associated with the $Ca^{40}(d,p)Ca^{41}$ reaction, but no evidence of higher-energy protons could be found. In order to improve counting rates the solid angle of observation was improved by using a single counter in spite of the background. The difficulty of a slightly varying background was overcome by keeping operating conditions extremely stable. Figure 2 shows the results of the most successful of approximately twenty-five attempts. The proton groups due to the reaction on Ca^{43} are only 20 percent above the background rate, but due to the stability of the background level their positions could be determined reasonably well.

The Q-value of 9.07 ± 0.07 Mev for the ground state peak is in fair agreement with the one calculated from mass values.[11] (See Table II.) Two excited states were observed at 1.15 and 2.28 Mev, in good agreement with (α,p) results. The cross section for this reaction for 90° observation was less by roughly a factor of fifty from that observed for $Ca^{40}(d,p)Ca^{41}$, considering ground states. Comparing first excited states the ratio of cross sections is a hundred. When comparing the yield to that obtained in the $K^{41}(\alpha,p)Ca^{44}$ reaction, the latter still has a cross section greater by a factor of seventy. Because of the nature of the experiment no reasonable estimates of absolute cross sections can be made. The relative cross sections are accurate within a factor of two.

The peak on the left of Fig. 2 is due to the $O^{16}(d,p)O^{17}$ reaction. Oxygen was present in all targets as a contaminant and the large yield of this reaction provided a convenient means of calibrating the beam energy.

V. $K^{41}(\alpha,p)Ca^{44}$

The yield from this reaction was perhaps a factor of five below that from K^{39}, but this was still adequate to measure the range of proton groups. A typical run, using a target containing 99 percent K^{41} and 1 percent K^{39} [10] is shown in Fig. 3. The angle of observation here

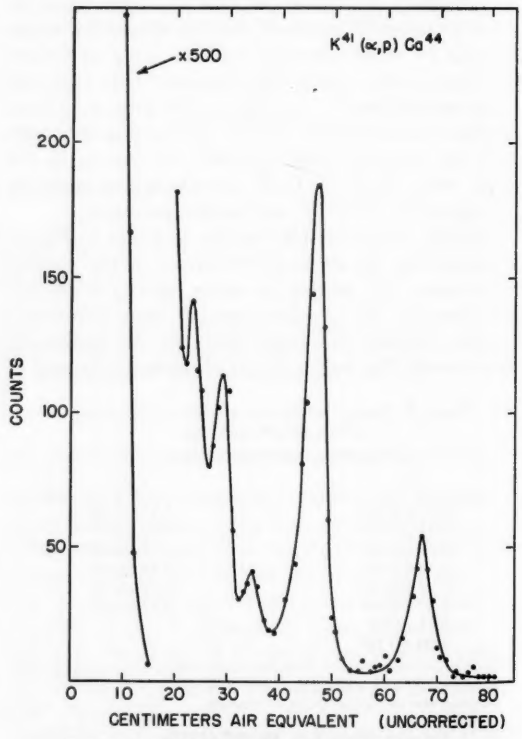


FIG. 3. A typical curve of the proton yield at 0° from $K^{41}(\alpha,p)Ca^{44}$.

[10] Separated isotopes were obtained from the Oak Ridge National Laboratory.

[11] Collins, Nier, and Johnson, Phys. Rev. 84, 717 (1951).

was 0°, with the peak on the left representing protons elastically scattered by the alpha beam from hydrogenous impurities in the target. The Q-value for the ground state was found to be 0.98±0.10 Mev. There is a small discrepancy between this value and the one calculated from mass values of Collins, Nier, and Johnson.[11] Earlier in these investigations it was thought that a level was detected beyond the one assigned to the ground state group, but further data showed that this was a spurious effect due to statistical fluctuations in background rates.

The excited states calculated are shown in Table I. Above the fourth excited state the level density was too high for the resolution of the equipment. The 1.92-Mev state was not seen with the (d,p) reaction, presumably because of its low yield. The gamma rays observed by Cohen, as mentioned in the previous section, could be assigned to the levels measured in this work, within the estimated experimental errors. The 1.13-Mev gamma ray was probably a first excited state to ground

TABLE II. Ground state Q values.

Reaction	Ground state Q (Mev) Measured in present work	Calculated from mass values[a]
$K^{39}(\alpha,p)Ca^{42}$	−0.19±0.07	0.34±0.10
$K^{41}(\alpha,p)Ca^{44}$	0.98±0.10	1.29±0.12
$Ca^{43}(d,p)Ca^{44}$	9.07±0.07	9.20±0.12

[a] See reference 11.

state transition, although the third to first excited state spacing of 1.15 Mev could perhaps also account for it. The 2.07- and 2.48-Mev gamma rays were probably due to levels of higher excitation not measured here.

VI. CONCLUSIONS

All available information on the level structure of Ca isotopes is shown in Fig. 4. The data on Ca^{40} were obtained by inelastic proton scattering by Braams.[12] The spin assignments in Ca^{41} are from Holt and Marsham's[13] work on angular distributions. Jeffries[14] determined the spin of Ca^{43} by magnetic moment measurements. The ground state and first excited

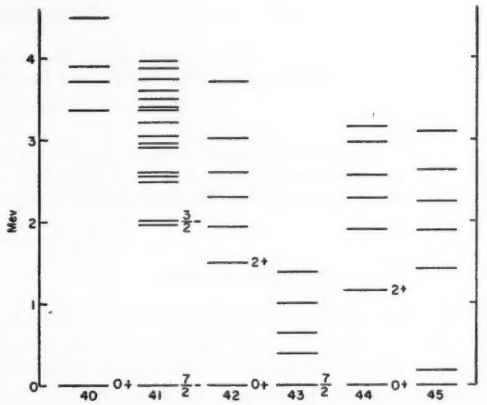


FIG. 4. Level structure in Ca isotopes. Levels in Ca^{42} and Ca^{44} are from the present work; the other isotopes were measured by Braams (see references 11 and 14). The first excited states decrease gradually above the doubly magic Ca^{40}.

state angular momentum assignments in Ca^{42} and Ca^{44} have been discussed in Secs. III and V. There is a surprising similarity between the excitation of the second to fifth states in Ca^{42} and Ca^{44}.[15] The level structures of Ca^{41}, Ca^{43}, and Ca^{45} are from work done by Braams[16] using magnetic analysis. The ground state angular momenta are in agreement with shell theory, which predicts that the independent particle is in an $f_{7/2}$ state between neutron number 20 and 28. The first excited state spacings show the predicted decrease with increasing neutron number above the closed shell at 20, if even-even and even-odd nuclei are viewed separately.

The anomalously low cross section for the $Ca^{43}(d,p)Ca^{44}$ reaction remains unexplained. The spin change involved is the same as that in the $Ca^{40}(d,p)Ca^{41}$ reaction. A shift in the angular distribution could account for only a very small fraction of the observed decrease in yield. Runs were also tried at 0° and 45° without success.

VII. ACKNOWLEDGMENTS

The author would like to express his appreciation to Professor Ernest C. Pollard for his suggestion of this problem and his advice in the course of this work.

[12] Braams, Bockelman, Browne, and Buechner, Phys. Rev. **91**, 474 (1953).
[13] J. R. Holt and T. N. Marsham, Proc. Phys. Soc. (London) **A66**, 565 (1953).
[14] C. D. Jeffries, Phys. Rev. **90**, 1130 (1953).
[15] D. Kurath, Phys. Rev. **91**, 1430 (1953).
[16] C. M. Braams, Phys. Rev. **95**, 763 (A) (1954); **95**, 650 (A) (1954).

Beta Spectrum of C^{14}

A. V. Pohm,* R. C. Waddell, J. P. Powers, and E. N. Jensen
Institute for Atomic Research and Department of Physics, Iowa State College, Ames, Iowa
(Received October 14, 1954)

A uniform thin source enriched with C^{14} was formed by vacuum evaporation of metallic lithium which was subsequently exposed to water vapor and carbon dioxide. The Kurie plot shows no deviation from linearity above 7 kev. The maximum energy of the beta spectrum was found to be 158.5±0.5 kev.

I. INTRODUCTION

THE beta spectrum of C^{14} has been studied carefully by many investigators,[1–10] and while general agreement exists on the end-point energy a question still remains as to its shape. Past studies by Cook, Langer, and Price,[2] Warshaw,[6] and Angus, Cockcroft, and Curran,[3] along with the recent results of Mize and Zaffarano,[9] all show deviations from a straight line Kurie plot. Zweifel,[11] by calling attention to second-order correction terms, hopes to show that the concave shape observed by Mize is consistent with beta-decay theory. In contradiction to the above results Feldman and Wu,[5] and Wu and Schwarzschild[8] have presented experimental evidence that the Kurie plot is straight to 25 kev.

A great difficulty in C^{14} studies has been the preparation of thin uniform sources. Many of the past investigations[2,4–6,9,10] have been with sources prepared by depositing suspensions or solutions containing C^{14} on films. These methods result in lumps, either as a result of original granules or crystallization in drying. Investigations by Lane[10] showed that carefully prepared sources of average density 0.03–0.05 mg/cm² had densities of approximately 1.5 mg/cm² in the individual grains. Wu and Schwarzschild[8] were able to obtain a thin (50-μg/cm²) source by vacuum evaporation of succinic acid onto a Formvar backing.

A method has been developed by the present authors which makes possible the vacuum evaporation of thin uniform sources of 6-μg/cm² surface density, containing enriched C^{14}. No protective coatings are necessary. Using thin uniform sources in an intermediate-image type spectrometer,[12] one obtains straight Kurie plots to 7 kev after correcting for the efficiency of the scintillation counter.

II. PROCEDURE

To obtain uniform thin C^{14} sources, the following procedure was used. Metallic lithium was vacuum-evaporated onto thin collodion films. Lithium was used because its low Z minimized scattering. While still in the vacuum, the films were placed in a small vacuum desiccator. After sealing the desiccator it was removed from the vacuum chamber and radioactive carbon dioxide and water vapor were admitted into it. The water vapor reacted with the metal forming lithium hydroxide and this in turn reacted with the carbon dioxide to form lithium carbonate. At the low pressures used, the complete reaction required a few hours. The low reaction rate prevented film breakage from the heat of reaction. Within experimental error, the completed sources were found to be uniform. Surface densities were measured with a microbalance.

To prevent possible source charging after they were placed in the spectrometer, the sources were immersed in a cloud of electrons every few minutes. The heated filament supplying the electrons was not left on while data were taken, because it increased the background count in the photomultiplier tube.

An estimate was made of the capacity of the sources and from their known activity it was possible to calculate the rate at which the potential of the sources would increase under the extreme condition that no leakage of charge occurred. The periods between turning on the filament were chosen so that even under this extreme condition, charging would be negligible.

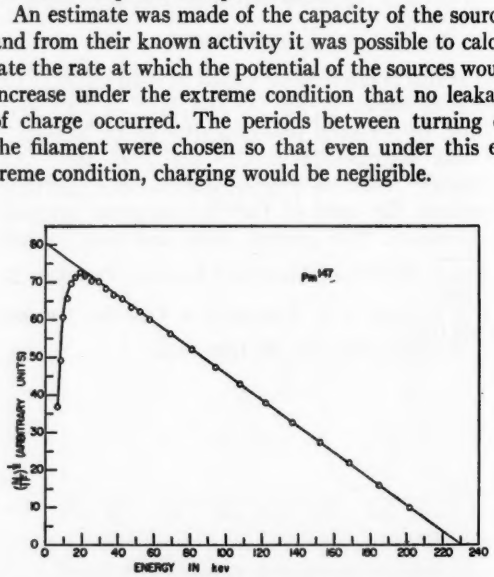


FIG. 1. Kurie plot of Pm^{147}. Surface density of 0.2 μg/cm² on a collodion backing of 4 μg/cm².

* Now at Engineering Research Associates, St. Paul, Minnesota.
[1] P. W. Levy, Phys. Rev. **72**, 248 (1947).
[2] Cook, Langer, and Price, Phys. Rev. **74**, 548 (1948).
[3] Angus, Cockcroft, and Curran, Phil. Mag. **40**, 522 (1949).
[4] P. W. Levy, Atomic Energy Commission Report AECU-173, March, 1949 (unpublished).
[5] C. Feldman and C. S. Wu, Phys. Rev. **75**, 1286 (1949).
[6] S. D. Warshaw, Phys. Rev. **80**, 111 (1950).
[7] G. H. Jenks and F. H. Sweeton, Phys. Rev. **86**, 803 (1952).
[8] C. S. Wu and A. Schwarzschild, Phys. Rev. **91**, 483 (1953).
[9] J. P. Mize and D. J. Zaffarano, Phys. Rev. **91**, 210 (1953).
[10] R. Lane, Doctoral Dissertation, Iowa State College, 1953 (unpublished).
[11] P. F. Zweifel, Phys. Rev. **95**, 112 (1954).
[12] Nichols, Pohm, Talboy, and Jensen, Atomic Energy Commission Report No. ISC-345, 1953 (to be published).

To check the linearity of the spectrometer and to correct for the loss of counts below 25 kev because of counter efficiency, an extremely thin Pm^{147} source was used. Langer, Motz, and Price[13] have shown that Pm^{147} has an allowed shape above 8 kev. The Pm^{147} source was evaporated onto a 4-$\mu g/cm^2$ collodion film in a vacuum. It had a surface density of less than 0.2 $\mu g/cm^2$. A screening correction was made on the Pm^{147} data by interpolation in the table of Reitz.[14] The screening correction for C^{14} was found negligible. All data were corrected for the six percent resolution of the spectrometer.

III. RESULTS

Figure 1, a Kurie plot of Pm^{147}, illustrates the linear characteristic of the spectrometer. This information was used in correcting the C^{14} data at low energies.

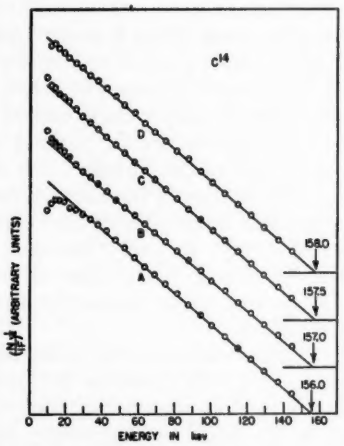


FIG. 2. Kurie plots of C^{14}. The source for curve A, 30–50 $\mu g/cm^2$ $BaCO_3$ on a backing of 4 $\mu g/cm^2$ was deposited from a suspension. The Li_2CO_3 sources for curves B, C, and D were vacuum evaporated on collodion backings 6 ± 2 $\mu g/cm^2$. Surface densities are: B, 40 ± 10 $\mu g/cm^2$; C, 25 ± 6 $\mu g/cm^2$; D, 10^{+5}_{-2} $\mu g/cm^2$.

The effects of variations in source thickness and preparation are shown in Fig. 2. Curve A was obtained by using a source with an average surface density of 30–50 $\mu g/cm^2$ prepared from a suspension by Lane[10] using ultrasonic agitation. Curves B, C, and D illustrate the effect of source thickness when prepared by the vacuum evaporation technique described above. Sources B, C, and D had surface densities of about 40, 25, and 10 $\mu g/cm^2$ respectively. The curves were corrected for

[13] Langer, Motz, and Price, Phys. Rev. 77, 798 (1950).
[14] J. R. Reitz, Phys. Rev. 77, 10 (1950).

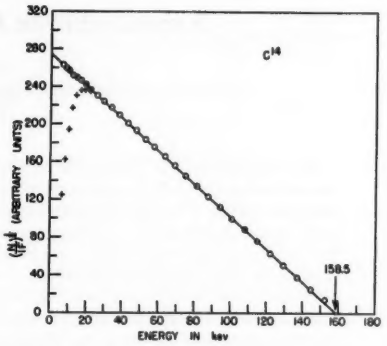


FIG. 3. Kurie plot of C^{14}. Vacuum evaporated source consisting of 6^{+4}_{-2} $\mu g/cm^2$ Li_2CO_3 on a collodion backing of 6 ± 2 $\mu g/cm^2$. Data corrected below 25 kev for detector efficiency. Crosses represent data before correcting for detector efficiency.

detector efficiency below 25 kev. It is interesting to note that although sources A and B had approximately the same average surface densities, their low-energy deviations differ markedly.

Figure 3 presents the results obtained by using a vacuum evaporated source having a surface density of (6^{+4}_{-2}) $\mu g/cm^2$ on a collodion film of (6 ± 2) $\mu g/cm^2$. The Kurie plot shows no deviation above 7 kev, within the probable error. This source gave a maximum counting rate of about 20 000 counts/min. Carbon enriched to about 40 percent C^{14} obtained from C. S. Wu was used in this source.

These data are consistent with the interpretation of an allowed transition for C^{14} according to Gamow-Teller selection rules with $\Delta I=1$, no.[8]

One notes that deviations from linearity as observed in curves B, C, and D of Fig. 2 and Fig. 3 are at the approximate energies predicted by the formula of Hamilton and Gross.[15]

It is observed that the end-point energy increases progressively with decreasing source thickness. A value of 158.5 ± 0.5 kev is obtained with the thinnest source. This is in general agreement with the results of other investigators.

IV. ACKNOWLEDGMENTS

The authors wish to express their gratitude to Dr. C. S. Wu for making available the enriched carbon used in some of the sources and to R. O. Lane, J. H. Talboy, Jr., and E. W. McMurry for suggestions and assistance during the experiment.

[15] D. R. Hamilton and L. Gross, Rev. Sci. Instr. 21, 912 (1950).

Experiments on the Direct Photonuclear Effect*

SVEN A. E. JOHANSSON †
Institute for Atomic Research and Physics Department, Iowa State College, Ames, Iowa
(Received June 23, 1954)

An investigation has been made of the high-energy protons and neutrons emitted in irradiation with bremsstrahlung of 65-Mev maximum energy. Protons above 14 Mev were measured. The energy limit for the neutrons varied. The angular distributions of protons from carbon, aluminum, nickel, and molybdenum could be fitted with curves of the form $a+(\sin\theta+b\sin\theta\cos\theta)^2$ with the maximum around 60°. The angular distributions of neutrons above 5 and 10 Mev could be fitted with curves of the form $a+b\sin^2\theta$. The yield of protons for eight elements from carbon to molybdenum was approximately proportional to Z. The yield of neutrons above 7.5 Mev for 19 elements from carbon to lead was approximately proportional to N. Excitation curves were measured for the high-energy protons from aluminum and phosphorus. They have a threshold at about 25 Mev and rise to about 45 Mev. From there on they are nearly constant. An attempt was made to detect proton-neutron coincidences from a carbon target. No true coincidences were found. It is concluded that the measurements are in good agreement with an independent particle model but that they agree only partly with a deuteron model.

INTRODUCTION

IT is well known that the emission of protons and neutrons in photonuclear reactions can be accounted for fairly well by an evaporation model. The nucleus is heated by the absorption of a photon and is de-excited by the evaporation of one or several particles according to the statistical theory.[1] Recent experiments[2–8] on the emission of photoprotons at low and medium x-ray energies have shown, however, that in addition to this process there exists a type of direct photoemission. The energy spectrum of the photoprotons shows an excess of high-energy protons compared to the spectrum expected from the statistical theory. There is also a departure from spherical symmetry in angular distribution. This process is relatively most important for medium and heavy elements where it gives an anomalously high ratio between the cross sections for (γ,p) and (γ,n) processes.[9] Work[10–15] with very high-energy x-rays at 200–300 Mev has shown the presence of high-energy protons emitted in a strongly forward direction. Very little is known about the emission of high-energy photoneutrons. Measurements[16–17] made with threshold detectors indicate an excess of neutrons at right angles to the beam.

To account for these effects Courant[18] and Jensen[19] have proposed a direct photonuclear effect. It is suggested that the photon is absorbed by only a small part of the nucleus, perhaps only a single nucleon. A proton or neutron may then be emitted without sharing its energy with the rest of the nucleus. There is therefore no statistical distribution of the excitation energy and the proton emission is much higher than expected from the statistical theory. In heavy elements, where the barrier greatly reduces the number of evaporation protons, this process might account for almost all of the emitted protons.

Levinger [20] has proposed another model for the high-energy photoeffect. He considers the photodisintegration of a quasi deuteron moving inside the nucleus. According to this model the high-energy protons and neutrons are emitted in pairs. The angular distribution of the particles is essentially the same as for the photodisintegration of the deuteron.

Comparing theory and experiment one encounters several difficulties. The experiments have been performed at different energies. The direct photoeffect is mixed with various other processes, different for different energies. With bremsstrahlung of 20–25 Mev maximum energy it is difficult to separate the directly emitted particles from the evaporation particles. With high-energy bremsstrahlung one might expect mesonic effects to play an important role, making the particle emission a fairly complicated process. Another difficulty is that most experiments deal with photoprotons; very little is known about the directly emitted photoneutrons.

The present work has been done using bremsstrahlung of 65-Mev maximum energy from the Iowa State College synchrotron. At this energy it is easy to separate the directly emitted particles from the statistically evaporated ones, yet the energy is sufficiently low so that

* Work performed in the Ames Laboratory of the U. S. Atomic Energy Commission.
† On leave from the University of Lund, Lund, Sweden.
[1] V. F. Weiskopf and D. H. Ewing, Phys. Rev. **57**, 472 (1940).
[2] Curtis, Hornbostel, Lee, and Salant, Phys. Rev. **77**, 290 (1950).
[3] B. C. Diven and G. M. Almy, Phys. Rev. **80**, 407 (1950).
[4] Mann, Halpern, and Rothman, Phys. Rev. **87**, 146 (1952).
[5] H. Hendel, Z. Physik **135**, 168 (1953).
[6] W. A. Butler and G. M. Almy, Phys. Rev. **91**, 58 (1953).
[7] M. E. Toms and W. E. Stephens, Phys. Rev. **92**, 362 (1953).
[8] M. M. Hoffman and A. G. W. Cameron, Phys. Rev. **92**, 1184 (1953).
[9] O. Hirzel and H. Wäffler, Helv. Phys. Acta **20**, 373 (1947).
[10] D. Walker, Phys. Rev. **81**, 643 (1951).
[11] C. Levinthal and A. Silverman, Phys. Rev. **82**, 822 (1951).
[12] J. C. Keck, Phys. Rev. **85**, 410 (1952).
[13] S. Kikuchi, Phys. Rev. **86**, 41 (1952).
[14] J. W. Rosengren and J. M. Dudley, Phys. Rev. **89**, 603 (1953).
[15] J. W. Weil and B. D. McDaniel, Phys. Rev. **92**, 391 (1953).
[16] H. L. Poss, Phys. Rev. **79**, 539 (1950).
[17] Demos, Fox, Halpern, and Koch, Phys. Rev. **86**, 605 (1952).
[18] E. D. Courant, Phys. Rev. **82**, 703 (1951).
[19] P. Jenson, Naturwiss. **35**, 190 (1948).
[20] J. S. Levinger, Phys. Rev. **84**, 43 (1951).

mesonic effects cannot occur. The aim of this work was to collect as much experimental information as possible about the high energy photoprotons and photoneutrons. The angular distribution has been measured both for protons and neutrons for various elements. The yield of high-energy protons and neutrons has been measured for different elements. Excitation curves for protons above a certain energy have been determined. Finally an experiment has been carried out in order to determine whether or not the protons and neutrons are emitted in coincidence.

EXPERIMENTAL PROCEDURE

General Description

The experimental arrangement for the measurements on protons is shown in Fig. 1. The bremsstrahlung beam was collimated with lead. The diameter of the beam was 3 cm at the position of the target. The photomultiplier detector could be rotated and be put in different angular positions with respect to the target. The target was similarly rotated, so that the effective thickness was always constant. Hence the energy loss of the protons in the target was independent of the counter position. When the apparatus is rotated, the dose received by the target changes because of the fact that the intensity decreases towards the edges of the beam. A correction has been applied for this effect; it is fairly small, however.

Essentially the same arrangement was used for the neutron measurements. The target was in this case a block of the material to be investigated. It remained in a fixed position while the counter was rotated. In both the proton and neutron experiments the counter was shielded from scattered radiation by a large pile of lead bricks.

The beam intensity was monitored by an ionization chamber placed to the side of the beam. No absolute determination of the dose was required but it was necessary to know the relative dose received by the target for a certain counting period. The machine proved so stable in operation that the most convenient way to arrange the experiments was to use a constant counting period and to keep the beam intensity constant by adjusting the machine to give a constant reading on the ionization chamber dosage ratemeter. The maximum deviations from the mean value on the meter were 1–2 percent. The variations of the mean values for different counting periods were less than 1 percent. The monitor readings were obtained on a recording instrument so that the constancy of the intensity could be checked and corrections applied. By activating foils of zinc and tantalum a further check on the constancy of the beam intensity was possible.

The counters will be described in greater detail in the following paragraphs. The pulses were taken out from cathode-followers and led through cables to the amplifiers in the control room. Care was taken to avoid any

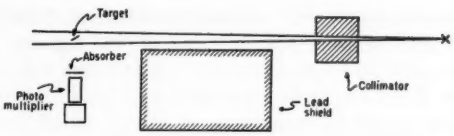


FIG. 1. Experimental arrangement for the proton measurements.

deformation of the pulses. They were analyzed by integral and single-channel discriminators and counted by conventional scalers. One of the major difficulties in these experiments was the piling up of the pulses. Although the mean counting rate was low, the counting rate during the time of a beam pulse was quite high. To keep the pile-up low, the pulses from the counters were made as short as possible. The necessity of reliable pulse height discrimination sets a lower limit for the pulse width. Another limiting factor is the time constant of the ZnS phosphor used in the counters. It is impossible to make the pulse length very much shorter than the time constant of the phosphor because this introduces great statistical fluctuations in the pulse height. A time constant of 0.75 μsec was found to be a suitable value for the pulses.

Another way to diminish the pile-up is to make the beam pulse duration as long as possible. No special pulse stretching device was used but the electrons were allowed to strike the synchrotron target near the maximum of the magnetic field. This gave a pulse length of about 100 μsec. Although this improved the situation, the pile-up was still a very serious limitation, especially in the proton experiments.

The Proton Counter

The proton counter consists of a thin, uniform layer of zinc sulfide on the top of a photomultiplier, type E.M.I. 6260. The grain size of the zinc sulfide (RCA 33Z-20-A) was reduced in the following way. The ZnS powder was suspended in absolute alcohol in a beaker and allowed to settle down partly. The part remaining in suspension was isolated rapidly from the rest by pouring the alcohol into another beaker, leaving the coarse grains on the bottom of the first beaker. This process was repeated several times. The fine-grained zinc sulfide obtained in this way was finally suspended in alcohol and allowed to settle completely down on the top of the multiplier. For this purpose a tube of plastic was tightly mounted on the top of the multiplier, to form a beaker for the suspension. Great care was taken to prevent convection currents in the liquid during the settling period. Finally most of the alcohol was taken away and the rest was allowed to evaporate. The resulting layer of zinc sulfide was very compact and durable. The thickness was 10 mg/cm^2. It was covered with a reflector of very thin aluminum foil. The photomultiplier was wrapped with black electrical tape and painted with Glyptal to make it light-tight.

In order to discriminate against the low-energy

evaporated protons, an aluminum absorber of thickness 274 mg/cm² was put in front of the counter. The distance from target to the zinc sulfide screen was 11 cm. The aluminum reflector was very thin (0.2 mg/cm²). These values correspond to an energy cutoff for the counter of 14 Mev.[21]

Figure 2 shows the pulse height spectrum obtained from an aluminum target of thickness 274 mg/cm². The steeply rising portion to the left is due to the piling up of pulses from the electron background. Fortunately the spectrum drops almost to zero before the proton pulses start to register. By setting an integral discriminator at a value slightly higher than the lowest point, one can completely discriminate against the electron pulses without losing more than a small fraction of the proton pulses.

It is possible to calculate the pulse height distribution if one knows the energy spectrum of the protons and the range-energy curve. The energy spectrum has been taken from the work of Hoffman and Cameron.[8] They found for the high-energy protons a differential spectrum of the form kE^{-n}. For aluminum bombarded with 65-Mev bremsstrahlung, n has the value 7. It turns out, however, that the results of the calculation are rather insensitive to the value of n. The calculated pulse height distribution has been normalized to give the best possible fit to the experimental points. The line in Fig. 2 shows the calculated curve. The agreement is very good, indicating that the counter works properly. The pulse height distribution has been calculated only for the lower part. The large pulses correspond to low-energy protons, and the range-energy curve is not known with sufficient accuracy in that energy range. It is, however, easy to calculate the number of pulses above a certain pulse height. A comparison with the experimental value gives a satisfactory agreement.

It has been assumed in the foregoing that all the heavily ionizing particles are protons. However, one might expect some deuterons. The absorption in the absorber and the air corresponds to a cut-off energy of 19 Mev for deuterons. It seems likely that the number of deuterons above 19 Mev is much less than the number of protons above 14 Mev. An experimental indication against the presence of deuterons is the good agreement between the experimental and the calculated pulse height distributions. If there were an appreciable number of deuterons present, they should cause a deformation of the pulse height distribution.

It is interesting to calculate the energy distribution of the photoprotons being registered by the counter. We assume that the efficiency of the counter is 100 percent, although this is not exactly true since some of the smallest proton pulses may fall below the discrimination level. They would correspond to high-energy protons. However, only a small part is lost in this way and our assumption is therefore approximately true. The protons are assumed to be emitted with an energy distribution of the form kE^{-n}. By taking the target thickness into account, the curve c in Fig. 3 is obtained. A value of $n=7$ was used in this calculation.

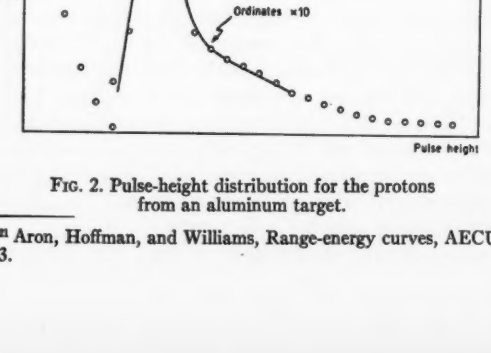


FIG. 2. Pulse-height distribution for the protons from an aluminum target.

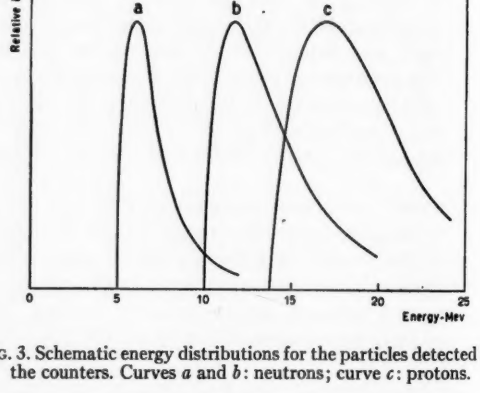


FIG. 3. Schematic energy distributions for the particles detected by the counters. Curves a and b: neutrons; curve c: protons.

The main features of the curve are the same for other reasonable values of n. The greater part of the protons registered by the counter falls in a fairly narrow energy range just above the threshold. Seventy-five percent of the protons have an energy between 14 and 21 Mev.

The Neutron Counter

The neutron scintillation counter is of a type devised by Hornyak.[22] The phosphor, however, was made in a different way. It consists of a suspension of zinc sulfide (R.C.A. 33Z-20-A) in plastic. The following procedure was used: The zinc sulfide powder was suspended in liquid plastic and a hardening catalyst was added. The liquid was stirred until it was very viscous, in order to prevent a nonuniform distribution of the zinc sulfide, and was then allowed to cure completely. Finally the plastic piece was machined to a cylinder, 35 mm in diameter and 15 mm thick, and was mounted on the top of a photomultiplier, type E.M.I. 6260. Silicone oil

[21] Aron, Hoffman, and Williams, Range-energy curves, AECU-663.

[22] W. F. Hornyak, Rev. Sci. Instr. 23, 264 (1952).

was used to make optical contact. An aluminum reflector was mounted around the plastic phosphor, and the whole counter was wrapped with black electrical tape and painted with Glyptal to become light-tight.

The basic process in this counter is the scattering of neutrons in the hydrogen of the plastic. The recoil protons are detected by the zinc sulfide. It has been shown[23] that the zinc sulfide alone detects neutrons by an (n,p) reaction in sulphur. If we know the cross section[24] of the (n,p) reaction and the amount of zinc sulfide in the phosphor (10 percent) it can be shown that the zinc sulfide itself contributes with less than 1 percent to the total detection efficiency in the present case.

The great advantage with a counter of this type is the low sensitivity to gamma rays and electrons. The intensity of the scattered gamma rays and electrons is so high, however, that even a counter of this type becomes overloaded. Therefore a lead filter 5 cm thick was used in front of the counter.

Figure 4 shows the pulse height distribution for the neutrons from an aluminum target. The upper part of the curve is linear but the counting rate increases more rapidly for small pulse heights. Scattering and other secondary effects will, of course, give an excess of low-energy neutrons. However, one might ascribe the high counting rate of small pulses to the presence of the low-energy evaporation neutrons. The upper part of the curve could then be due to high-energy, directly emitted neutrons. The break in the curve comes at about 4 Mev which is in reasonable agreement with the interpretation given here.

In order to calibrate the counter a polonium-beryllium source was used. The energy distribution of the neutrons from such a source is well known.[25] The corresponding energy distribution for the protons was then calculated. It was assumed that the pulse height is proportional to the proton energy and that the resolution of the counter is fairly good. It is rather likely that both these assumptions are wrong because of the inhomogeneous character of the phosphor. However, the pulse height distribution calculated in this way agrees in shape fairly well with the one obtained experimentally with the Po-Be source. This fact makes it possible to find the end point of the pulse distribution by extrapolation to zero counting rate. The pulse height value so obtained corresponds to the energy value for the end point of the Po-Be neutron spectrum. There is, of course, a considerable uncertainty in this calibration. A conservative estimate of the error is ± 10 percent. Furthermore, only one calibration point is used. The calibration might be nonlinear, especially at low energies. However, for the purpose of this work it is not necessary to know exact energy values. It is sufficient to be able to discriminate against the low-energy evaporation neutrons.

[23] G. R. Keepin, Rev. Sci. Instr. **25**, 30 (1954).
[24] R. K. Adair, Revs. Modern Phys. **22**, 249 (1950).
[25] B. G. Whitmore and W. B. Baker, Phys. Rev. **78**, 799 (1950).

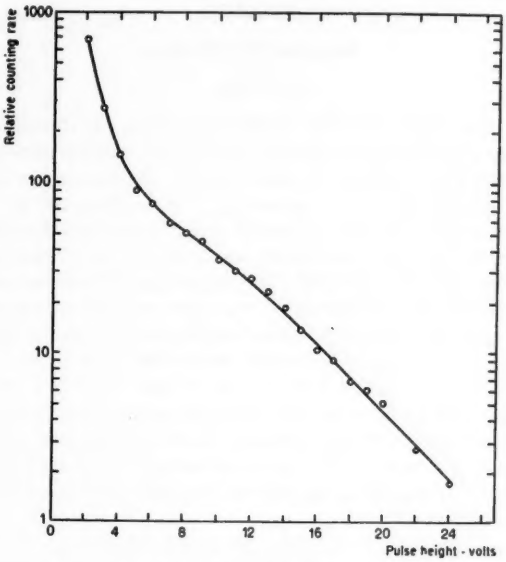


FIG. 4. Pulse-height distribution for the neutrons from an aluminum target.

An estimate can be made about the energy distribution of the neutrons detected by the counter for various discriminator settings. The primary energy distribution of the neutrons is assumed, in analogy with the proton case, to be of the form kE^{-n}. The value of n is not known. It has therefore been assumed tentatively that n has the same value for neutrons as for protons of the same energy. This gives a value of n of about 5. However, the exact form of the neutron spectrum is less important for this approximate calculation. It is sufficient to know that the number of neutrons falls off rapidly with the energy. The assumed energy spectrum has to be corrected for the absorption in the lead filter. Multiplication by the sensitivity of the counter as a function of the energy then gives the energy distribution of the counted neutrons. The curves a and b in Fig. 3 show the spectra for discriminator settings corresponding to thresholds at 5 and 10 Mev, respectively. It can be seen that the main part of the neutrons fall in a fairly narrow range just above the threshold.

The curves a and b in Fig. 3 have a sharp lower limit corresponding to a well-defined threshold for the counter. However, this requires a perfect phosphor with very high energy resolution. Actually the counter has a poor resolution. Recoil protons of a certain energy give pulses with a wide variation in pulse height. Some of them may be considerably higher than the mean value. Therefore, neutrons with an energy below the threshold value have a certain chance to be counted. It means that the sensitivity curve of the counter has a tail at the low-energy side. The importance of this fact will be discussed later.

RESULTS

Angular Distribution

Protons

The main difficulty in this experiment was the piling up of pulses in the counter, because of scattered gamma rays and electrons. It could be kept sufficiently low by lowering the beam intensity. The requirement of a reasonable number of counts sets a lower limit for the intensity. The synchrotron was operated fairly close to this limit in order that the measurements could not be affected by the pile-up. It was regarded better to have somewhat greater statistical errors than to run the risk of introducing systematical errors. Before each run a pulse distribution like the one shown in Fig. 2 was recorded to check the complete separation of the protons from the electron background. To get a further check the pulses were displayed on an oscilloscope, type DuMont 404A. Any piling up showed up very clearly on the oscilloscope picture.

The pile-up limits this experiment in two respects. It proved to be impossible to investigate the high-energy protons from the heavy elements. The gamma-ray and electron background increases more rapidly with the mass number than does the number of protons. For mass numbers higher than 100, the background was so heavy that it was impossible to operate the counter properly. Another limitation is that most of the gamma rays and electrons are scattered into the forward direction. Therefore it was impossible even for light elements to measure for smaller angles than 45°.

The following four elements were investigated: carbon, aluminum, nickel, and molybdenum. The targets were 25 mm wide and 60 mm long. Values for the target thickness are listed in Table I. The targets were mounted 11 cm from the zinc sulfide screen and followed the rotation of the counter. The plane of the target made an angle of 70° with the axis of the counter.

Figure 5 shows the angular distributions for carbon, aluminum, nickel, and molybdenum. The distributions have been measured several times and the agreement has been within the statistical errors. The experimental points have been fitted with curves of the form $a+(\sin\theta+b\sin\theta\cos\theta)^2$. The values on the constants are listed in Table I. The form of the curves indicates that the forward asymmetry is caused by interference between

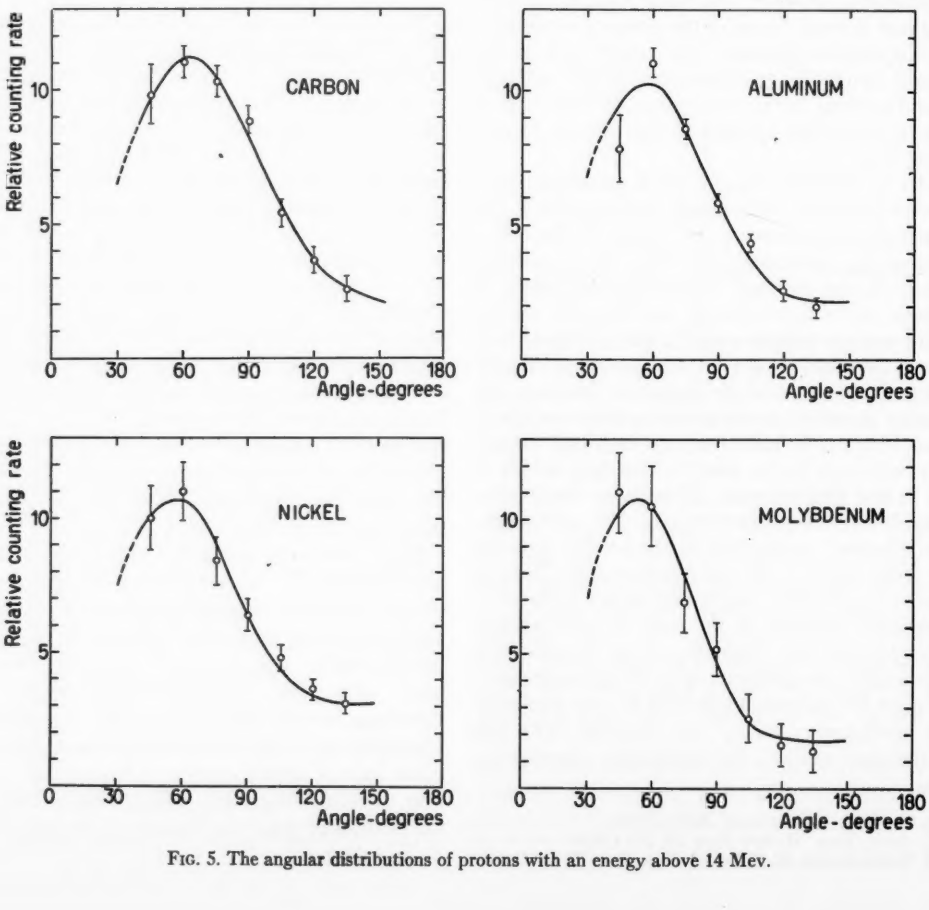


FIG. 5. The angular distributions of protons with an energy above 14 Mev.

TABLE I. Target thickness and the constants a and b in the angular distribution curve $a+(\sin\theta+b\sin\theta\cos\theta)^2$.

Element	Target thickness mg/cm²	a	b
Carbon	182	0.32	0.80
Aluminum	274	0.58	1.35
Nickel	352	0.94	1.45
Molybdenum	295	0.62	2.00

electric dipole and electric quadrupole photon absorption. The constant b increases with the atomic number, giving an increased forward shift for the heavier nuclei. This means that the quadrupole absorption increases with the atomic number.

It is interesting to compare these results with experiments at other energies. Mann, Halpern, and Rothman[4] measured the angular distribution of photoprotons using bremsstrahlung of 23-Mev maximum energy. They got curves peaked around 70°. Hendel[5] used bremsstrahlung of 150-Mev maximum energy and got curves peaked around 50°. Hence the maximum of the angular distribution shifts forward with increasing photon energy.

The constant a is a measure of the isotropic part of the distribution. The isotropic part varies from about 30 percent in carbon to about 50 percent in nickel. One should expect in the direct photoprocess that a certain part of the emitted protons would make collisions with the other nucleons without losing more than a small part of their energy before escaping. Such a process would give an isotropic part in the angular distribution. The nuclei become more transparent with increasing proton energy and therefore the isotropic part should become less. This is actually the case. Angular distributions for 30- and 50-Mev protons[5] show a considerably smaller isotropic part than the present experiment. It is impossible to decide whether this internal scattering can account for the whole isotropic part. More has to be known about the mechanism of the direct photoemission before this question can be solved. It is interesting to note that the calculations of Courant based on the independent-particle model give a considerable isotropic part. The size of it depends on the quantum state of the emitted nucleon in the nucleus.

Neutrons

The targets in this experiment were blocks of beryllium, aluminum, tantalum, and lead. They were kept in a fixed position. The neutron counter was set in different angular positions around the target. The target-counter distance was 17 cm except for the two smallest and two largest angles, where it was 21 cm. In a separate experiment it was confirmed that the counting rate varied with the inverse square of the distance. The counting rates at distance 21 cm could then be converted to counting rates at 17 cm.

The elastic scattering of the neutrons presents a difficult problem in these experiments. It takes place within the target, in the lead absorber, and also in the material around the counter (lead shield, support). The scattering in the target can be estimated from the size and shape of the target. It turns out to be small. The scattering in the absorber and the surroundings is very difficult to calculate. A very rough estimate shows that it too is small. Because of the complicated nature of the calculations, no attempt has been made quantitatively to correct the measured values for the scattering. It seems certain, however, that it cannot have any essential influence on the results.

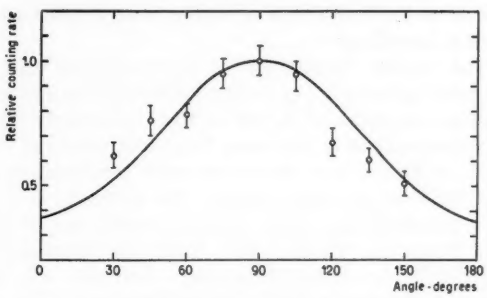


FIG. 6. The angular distribution of the neutrons from beryllium. Counter threshold at 10 Mev.

The synchrotron was run at constant intensity during this experiment. The intensity was monitored by an ionization chamber as described above. The counter was set at different angular positions and the counting rate was determined. The counting rate without target was then determined under exactly the same conditions and this background was subtracted to get the true counting rate. The background was almost constant for all counter positions (about 20 percent of the counting rate at 90°). The pulse height spectrum falls off very rapidly with increasing pulse height. Therefore the stability of the apparatus is very essential. In order to make sure that the results were not affected by any drift in the electronic circuits the following procedure was followed. The counting rate was first determined at 90°. After a number of readings had been taken in other positions, it was again measured at 90°. If the two counting rates did not agree within the combined standard errors, the whole series was discarded. In most cases, however, the stability proved to be sufficient. The curves were quite reproducible.

Figure 6 shows the angular distribution of the photoneutrons from beryllium. The threshold of the neutron counter was set at 10 Mev. The experimental points have been fitted with a curve of the form $a+b\sin^2\theta$. The curves for thresholds at 5 and 2.5 Mev are very similar to the one shown in Fig. 6. One might expect them to have a greater isotropic component due to the low-energy evaporation neutrons. The reason that there is no such increase of the isotropic part is probably that

the statistical model does not apply to very light nuclei such as beryllium.

The angular distribution of the aluminum photoneutrons is shown in Fig. 7. The threshold of the neutron counter was set at 2.5, 5, and 10 Mev, respectively. The experimental points have been fitted with curves of the form $a+b\sin^2\theta$. The two curves with thresholds at 5 and 10 Mev are quite similar. The curve for a 2.5-Mev threshold has a much higher isotropic component. The reason is probably that when the threshold is sufficiently low, the evaporation neutrons will be counted, thereby increasing the isotropic part.

Figure 8 shows the angular distribution of the photoneutrons from tantalum. Runs were taken with thresholds both at 5 and at 10 Mev. The experimental points could again be fitted with curves of the form $a+b\sin^2\theta$. The constants are very nearly the same as for the corresponding aluminum curves. The two points at 30° and 45° in the 5-Mev curve deviate considerably from the curve. This is attributed to the scattered gamma rays, which increase the counting rate by piling up. This effect is most likely to occur for heavy elements and in the forward direction. If the threshold is lowered to 2.5 Mev, the whole forward part of the curve is deformed giving a steady increase in counting rate from 150° to 30°.

The angular distribution for the photoneutrons from lead is shown in Fig. 9. The threshold was 5 and 10 Mev. Also in this case the experimental points can be fitted with curves of the form $a+b\sin^2\theta$. The constants have

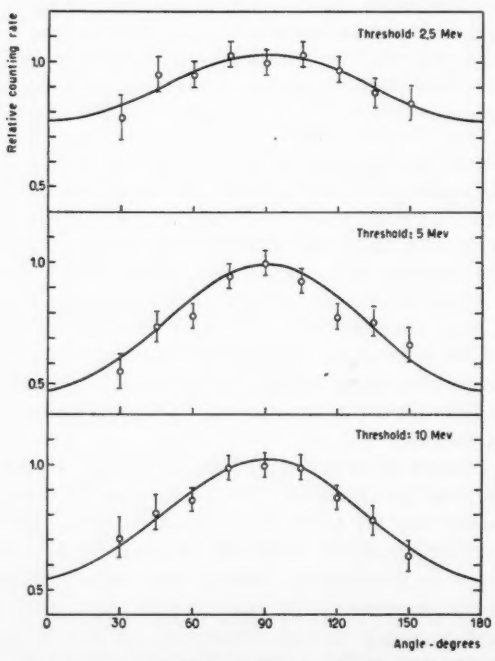


FIG. 7. The angular distributions of the neutrons from aluminum. Counter thresholds at 2.5, 5, and 10 Mev.

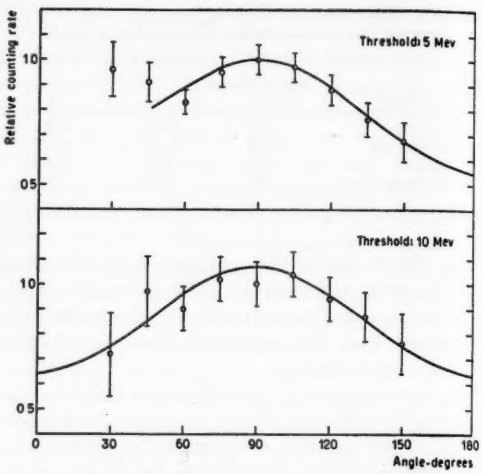


FIG. 8. The angular distributions of the neutrons from tantalum. Counter thresholds at 5 and 10 Mev.

about the same values as for aluminum and tantalum.

The isotropic part has almost the same relative intensity for all the 5- and 10-Mev curves except the one for beryllium. About 60 percent of the neutrons belong to the isotropic component. As discussed above only a small part of it can be attributed to scattering of the neutrons. An effect which might be of greater importance is the lack of an absolutely sharp threshold for the neutron counter. The curves in Fig. 3 have been obtained under very simplified assumptions. It is possible that the sensitivity curve has a tail at the low-energy side, in which case some of the evaporation neutrons might be counted. Even if the sensitivity is very low, the number of evaporation neutrons is so high that it might give a considerable contribution to the isotropic component. If this is true, however, one should expect the isotropic part to be relatively more important with the threshold set at 5 Mev than at 10 Mev. This is not the case. Another indication against this explanation is that Poss,[16] using aluminum as threshold detector, found about the same ratio for the counting rates at 90° and 0° as obtained in the present work. A threshold detector, of course, has a sharp energy cutoff (in the case of aluminum, at 4.6 Mev). Therefore it seems likely that the greater part of the isotropic distribution is real and not due to scattering or instrumental effects.

One might expect a certain fraction of the neutrons to collide with the other nucleons before leaving the nucleus. If the energy loss is small they may still be detected by the counter. This effect gives a contribution to the isotropic component. The same situation for the protons has already been discussed, and the difficulties are the same in this case. A quantitative calculation of the internal scattering is not feasible and therefore it is impossible to know if it can account for the whole isotropic component. It should be noted, however, that in this work the mean energy of the neutrons is lower

than that of the protons. The internal scattering should therefore be of greater importance for the neutrons.

The Yield

Protons

The targets in this experiment were cylindrical with a 44-mm diameter. They were made with varying thickness corresponding to the range of 14-Mev protons in the respective material. The cylindrical plates were mounted in the synchrotron beam so as to make an angle of 15° with the beam. The support was made of plastic film and thin aluminum foil. The counter was set in the 105° position.

The intensity of the synchrotron was kept constant as described above. The maximum bremsstrahlung energy was 65 Mev. The counting rate was determined for each target. The background without target was also determined and subtracted to get the true counting rate. The stability of the apparatus was checked by measuring the counting rate for a copper target several times during the run. The variations were within the statistical errors.

The yield was determined for eight elements: C, Mg, Al, P, S, Ni, Cu, and Mo. The relative yield per mole is plotted against the atomic number Z in Fig. 10. The experimental points fall very well along a straight line. The slope of the line is 0.92. Hence the yield of protons above 14 Mev is approximately proportional to Z.

Neutrons

The neutron yield was determined for 19 elements from carbon to lead. The targets were cylindrical, 44 mm in diameter and of weight 60 grams. They were mounted with the flat side perpendicular to the beam. The support was made of thin aluminum foil and plastic film. The counter was placed 10 cm from the target and in a direction making an angle of 105° with the beam. The bremsstrahlung is absorbed to some extent in the target so that the intensity is not constant over the volume of the target. This is especially true for the heavy elements. The thickness of the target was kept so low, however, that this effect was negligible in this experiment.

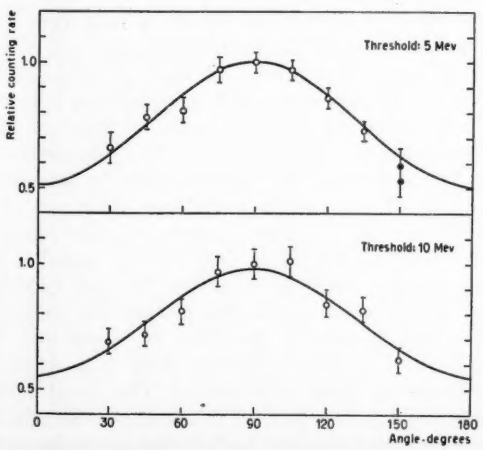


FIG. 9. The angular distributions of the neutrons from lead. Counter thresholds at 5 and 10 Mev.

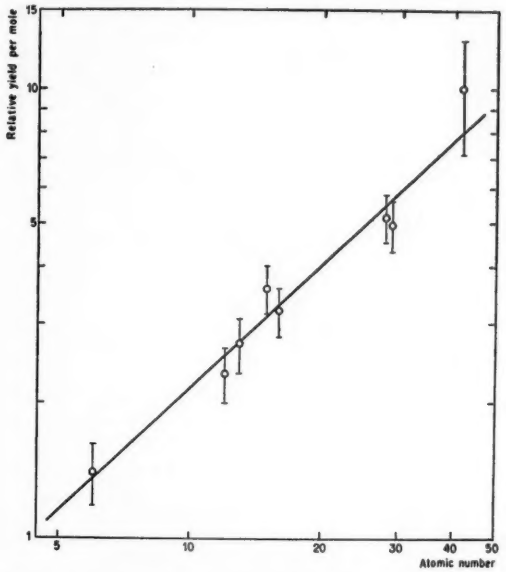


FIG. 10. The relative yield per mole for protons above 14 Mev as a function of the atomic number.

The threshold of the neutron counter was set at 7.5 Mev. The synchrotron was run at 65 Mev. The experiment was performed in the same way as for the protons. The beam intensity was kept constant and the counting rate determined with and without target. The stability of the apparatus was checked by measuring the counting rate for the aluminum target several times during the run.

The relative yield per mole is plotted in Fig. 11 as a function of the number of neutrons, N. The experimental points fall nicely along a straight line with a slope of 1.05. Hence the yield is nearly proportional to N.

Excitation Curves

The energy distribution of the protons detected by the counter is shown schematically in Fig. 3(c). The greater part of them fall in a fairly narrow range just above the threshold. The distribution is approximately symmetrical with a mean energy of 17.5 Mev and a half-value of 6 Mev. It would be very interesting to know the energy distribution of the photons which cause the emission of these protons. It means that we have to determine the yield of the high-energy protons as a function of the maximum energy of the bremsstrahlung. The dose received by the target was determined by

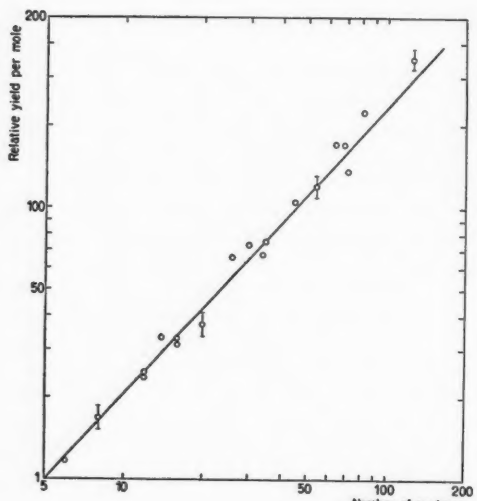


FIG. 11. The relative yield per mole for neutrons above 7.5 Mev as a function of the neutron number.

simultaneous activation of monitor foils. The following procedure was used: The synchrotron was run at a certain energy and with constant intensity. The number of high-energy protons from a thin target was counted for a certain length of time. Foils of zinc or tantalum were activated in the beam during the counting time. The activity of the foils was determined by a thin-walled Geiger-Müller tube under exactly the same conditions every time. Measurements were also made without target to determine the background. The counting rates and activities were determined for a number of different synchrotron energies.

The piling up of electron pulses was a serious limitation in this experiment as well as in all the other proton experiments. The beam intensity was kept very low to make absolutely sure that no pile-up took place. This, of course, made the statistics rather poor, but the results are sufficiently accurate to give the main features of the excitation curves.

Figure 12 shows the excitation curves for aluminum and phosphorus. The ratio between the number of high-energy protons and the activity of the monitor foil is plotted against the synchrotron energy. The monitor foils were in this case made of zinc. The curves are very similar. Starting from about 25 Mev, they rise to about 45 Mev; from there on they are fairly constant. It is rather difficult to make a detailed analysis to find the cross-section curves because of the considerable uncertainty in the activation curves. It is easy to see, however, that the greater part of the cross-section curves falls between 30 and 40 Mev and that they have a maximum around 35 Mev. Hence one can conclude that the protons with a mean energy of 17.5 Mev are mainly produced by photons in the energy range from 30 to 40 Mev.

Coincidence Experiment

For the high-energy photoeffect Levinger[20] has proposed a deuteron model. According to this model the photon interacts with a quasi deuteron inside the nucleus. The high-energy protons and neutrons are therefore emitted in pairs. A coincidence experiment was performed in order to see whether the directly emitted protons and neutrons could be produced by a process of this type also at lower energies.

A Lucite target 182 mg/cm² thick was mounted between the proton and neutron counter. The counters were placed in opposite directions at right angles to the beam. They were operated in the way described above. The threshold of the proton counter was 14 Mev as before and the discriminator of the neutron counter was set at 10 Mev. The output pulses from the discriminators were shaped and led to a coincidence circuit. The resolving time was determined from the width of the incoming pulses to be about 0.75 μsec. The counting rates in the two counters and the number of coincidence pulses were recorded.

With this experimental arrangement a certain number of coincidence pulses was obtained. The neutron counter was then moved 20 cm along the beam so that proton-neutron coincidences no longer should be recorded. Coincidence pulses were still obtained. There was no process which could give rise to true coincidences. Hence all these pulses must be accidental. A calculation using the counting rates, the length of the beam pulses, and the resolving time of the coincidence circuit gave a rate of coincidence pulses in reasonable agreement with the observed one. It was then possible, since the counting rates were known, to calculate the number of accidental coincidences in the first experiment. It turned out that the calculated number of accidental coincidences agreed within the statistical fluctuations with the observed number of coincidences. Hence no true coincidences between protons and neutrons were observed in this experiment.

This does not exclude the possibility that there is a certain number of proton-neutron pairs emitted. However, an upper limit can be set from the information obtained in the experiment. It is then necessary to know the solid angles, the sensitivity of the counters and the angular correlation in the emission of a proton-neutron pair. The solid angles are obtained from the geometry of the experiment and the sensitivities have been calculated. Nothing is known about the angular correlation. Therefore the upper limit for the emission of proton-neutron pairs is calculated for two extreme cases. If it is assumed that the protons and neutrons are emitted in opposite directions the upper limit is 3 percent. If it is assumed that there is no angular correlation the limit is 50 percent. The actual value should be somewhere between these two extreme values. This experiment therefore leads to the conclusion that the number of proton-neutron pairs is less than the number of single protons.

DISCUSSION

The experimental material presented here makes it possible to draw some conclusions about the direct photoeffect. We will compare the experiments with the theoretical calculations of Courant[18] and Levinger.[20]

Courant's calculations are based on an independent-particle model. The nucleons are assumed to move independently in a square well potential. The cross section for the direct photoeffect is equal to the sum of the cross sections for all the individual nucelons, only dipole photon absorption being taken into account. The angular distribution of the emitted particles is of the form $a+b\sin^2\theta$ for both protons and neutrons. This is exactly the distribution obtained for the high-energy neutrons, but it is in definite disagreement with the experimental angular distributions for the protons. The reason for this disagreement might be that only dipole photon absorption has been taken into account. A calculation[26] with electric quadrupole absorption included gives the following results: For neutrons there is no change because the effective charge of a neutron is zero in the quadrupole case. For protons, however, the electric quadrupole absorption gives a forward shift of the peak due to interference. This is in very good agreement with the experimental results.

According to the theory the yield of high-energy protons and neutrons varies roughly as $Z^{\frac{1}{3}}$. The disagreement with the experiment is perhaps not too serious. The experiment gives the yield of the protons and neutrons above a certain energy. Without knowing the energy distributions one cannot calculate the total yield of the directly emitted particles. A detailed comparison with the theory is, therefore, rather difficult in this case. For the same reason it is impossible to find an absolute value of the cross section for comparison with the theory.

It is easy to calculate approximately the excitation curves for proton production on the basis of an independent-particle model. Assuming a Fermi distribution inside the nucleus we find a mean binding energy of about 18 Mev for aluminum and phosphorus. The protons detected by the counter have a mean energy of 17.5 Mev. We therefore expect the cross section curves to have a maximum at about 35 Mev. This is in very good agreement with the excitation curves in Fig. 12.

The quasi-deuteron model of Levinger is intended to apply to energies above 150 Mev. It would be rather surprising to find it applicable also to the energy range investigated in this work. It might, however, still be of some interest to compare the experimental results with the predictions of the deuteron model. According to this model the angular distributions for protons and neutrons are different because of the interference between electrical dipole and electrical quadrupole terms. In the center-of-mass system the protons are emitted predominantly in the forward direction and the neutrons in the backward direction. In the laboratory system the photon momentum shifts both the proton and neutron distributions in the forward direction. The backward shift of the neutron distribution due to the interference is nearly cancelled by the forward shift due to the photon momentum. We therefore have good agreement between the experimental results and the deuteron model in this respect.

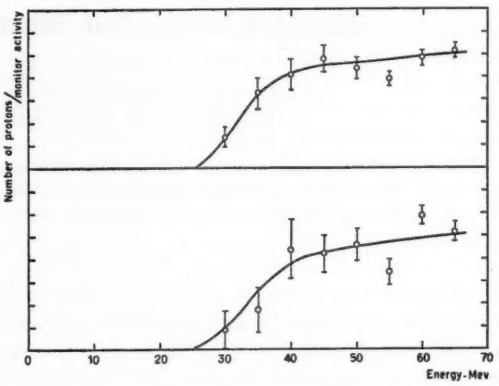


FIG. 12. The ratio between the number of high-energy protons and the activity of the monitor foil as a function of the sychrotron energy. The top curve is for aluminum and the bottom curve is for phosphorus.

According to the deuteron model, the cross section of the direct photoeffect is proportional to NZ/A, or nearly proportional to N and Z. There is good agreement with the experimental results also in this respect, though a literal deuteron model appears to be in very definite disagreement with the experimental excitation curves. One would expect the cross-section curves for the high-energy protons to have a threshold at about 45 Mev and a maximum at about 70 Mev. The excitation curve would then show a steady rise from the threshold at 45 Mev up to about 100 Mev. This is in complete disagreement with the experimental curves.

The coincidence experiment showed that only a rather small number of proton-neutron pairs are emitted. This is also in disagreement with the literal deuteron model. However, the experimental results do not exclude the possibility that proton-neutron pairs are formed inside the nucleus but that for some reason only one of the particles escapes without appreciable energy loss.

The conclusion of this comparison is that the independent-particle model, modified to take quadrupole absorption into account, is in reasonable agreement with the experimental results. The literal deuteron model, on the other hand, agrees with some of the experimental results but is in serious disagreement with others.

The author wishes to thank Dr. D. J. Zaffarano for his invaluable aid during the course of this research. The continued interest and encouragement of Dr. L. J. Laslett is gratefully acknowledged.

[26] The author is indebted to Dr. G. Källén for the calculation and for interesting discussions on this subject.

Effects of Isotopic Spin Selection Rules on Photonuclear Yields

H. MORINAGA*
Department of Physics and Institute for Atomic Research, Iowa State College, Ames, Iowa
(Received July 30, 1954; revised manuscript received September 29, 1954)

The effects of isotopic spin selection rules were taken into consideration in order to explain the observed high ratio of (γ,p) to (γ,n) cross sections. The recent experimental evidence on the separate giant resonances for (γ,p) and (γ,n) cross sections seems to support this viewpoint.

IT is known that with light nuclei ($Z \leq 20$) the (γ,p) reaction cross sections are about the same order of magnitude or sometimes even larger than the (γ,n) reaction cross sections for the same Z.[1-4] (See Table I.) For example, the $Al^{27}(\gamma,p)$ cross section is about the same order of magnitude as the $Al^{27}(\gamma,n)$ cross section. This might be explained by the fact that the (γ,p) threshold is about 5 Mev lower than the (γ,n) threshold, and therefore the suppression of proton emission by the Coulomb barrier is counter balanced by higher available energies for protons than neutrons and, also, possibly by there being more levels accessible for protons.

This argument does not hold when $Mg^{25}(\gamma,p)$ or $Mg^{26}(\gamma,p)$ yields are discussed. In these cases neutron thresholds are several Mev lower than the proton thresholds, and this might be expected to favor neutron emission strongly over proton emission. Experiments show, however, that the (γ,p) cross section appears to have the same order of magnitude as (γ,n) for these and other neighboring nuclides.

Recently it was shown by several authors that the isotopic spin selection rules hold with good approximation in nuclear reactions,[5] and especially it was shown that this is also true with photonuclear reactions in light nuclei involving photons in the range of electric dipole absorption.[6] Therefore, it seems to be necessary to take account of this effect when the magnitude of the photonuclear yield from light nuclei is estimated.

As an example, let us consider the case of the photonuclear reaction with Mg^{25} (Fig. 1).

The ground state of Mg^{25} has $T=\frac{1}{2}$, and by dipole absorption of γ rays the state with either $T=\frac{3}{2}$ or $T=\frac{1}{2}$ will be excited, since the selection rules allow $\Delta T=0$, 1 for dipole absorptions by nuclei for which $A \neq 2Z$.

The $T=\frac{3}{2}$ states can decay into either $T=1$ or $T=2$ states of neighboring nuclides, i.e., Na^{24} or Mg^{24} by proton or neutron emission, respectively. The $T=\frac{1}{2}$ states can decay into $T=0$ or $T=1$ states by nucleon emission.

The lowest $T=1$ state in Na^{24} is its ground state, and the lowest $T=1$ state in Mg^{24} lies about 9.4 Mev higher than its $T=0$ ground state. Above this lowest $T=1$ state the level scheme of Mg^{24} with $T \geq 1$ is identical with that of Na^{24} above its ground state, except for the Coulomb energy displacement of about 5 Mev.

Hence the decay from $T=\frac{3}{2}$ states in Mg^{25} is symmetric with regard to proton and neutron emission except for the difference in available energy for reaching the corresponding states in the two resulting nuclides, the suppressing of protons by the Coulomb barrier, and a factor $\frac{1}{2}$ for protons due to the coupling factor of isotopic spin. The first of these corrections enhances proton emission probability and the others suppress it. The situation is rather similar to the case of the (γ,p) and (γ,n) reactions in Mg^{24}. Here, product nuclides are mirror pairs whose level schemes are identical with each other and the decay of the photon-excited states (mostly $T=1$) by proton and neutron emission is symmetric in the sense mentioned above.

There are some reasons to suspect that proton emission can be more favored than neutron emission from $T=\frac{3}{2}$ states. The reason is as follows: The decay of the $T=\frac{3}{2}$ states in Mg^{25} is rather similar to that of the $T=1$ states in Mg^{24}, since the (γ,p) thresholds are about the same, and the (γ,n) threshold in Mg^{24} and the (γ,n) threshold from the $T=\frac{3}{2}$ states in Mg^{25} are about the same. Also the energy value of the (γ,n) peak of Mg^{24} is about the same as that of the (γ,p) peak of Mg^{25}. The (γ,p) reaction of Mg^{24} has not been measured, but guessing from the fact that the (γ,p) cross section on C^{12} is greater than the (γ,n) cross section on the same nuclide,[7] and also from the small magnitude of (γ,n) reaction on Mg^{24}, one may assume that the (γ,p) cross section probably exceeds the (γ,n) cross section with Mg^{24}, too. So, from the $T=\frac{3}{2}$ states of Mg^{25}, the proton emission can exceed the neutron emission.

From $T=\frac{1}{2}$ states in Mg^{25} neutrons compete strongly with protons, since by neutron emission the $T=0$ states in Mg^{24} can be attained. The decay into $T=1$ states of the product nuclides are symmetric in regard to proton and neutron emission but these transitions must compete with more favorable neutron emission which leads to the $T=0$ states, since the latter has a lower threshold than proton emission by about 5 Mev.

The foregoing discussion may be crudely summarized

* On leave from the University of Tokyo, Tokyo, Japan. Now at Purdue University, Lafayette, Indiana.

[1] O. Hirzel and H. Wäffler, Helv. Phys. Acta **20**, 373 (1947).
[2] L. Katz and A. G. W. Cameron, Phys. Rev. **84**, 1115 (1951).
[3] L. S. Edwards and F. A. MacMillan, Phys. Rev. **87**, 377 (1952).
[4] Katz, Haslam, Goldenberg, and Taylor, Can. J. Phys. **32**, 580 (1954); McPherson, Pederson, and Katz, Can. J. Phys. **32**, 593 (1954).
[5] For example: R. K. Adair, Phys. Rev. **87**, 1041 (1952).
[6] M. Gell-Mann and V. L. Telegdi, Phys. Rev. **92**, 169 (1953).
[7] J. Halpern and A. K. Mann, Phys. Rev. **83**, 370 (1951).

by stating that the $T=\frac{3}{2}$ states contribute principally to proton emission and the $T=\frac{1}{2}$ states give mostly neutrons. Thus the (γ,p) to (γ,n) ratio depends on the ratio of the probability of formation of $T=\frac{3}{2}$ states and $T=\frac{1}{2}$ states, rather than solely on the values of proton and neutron thresholds, as is assumed in the statistical model.

It is interesting to look at the recent measurements of (γ,p) and (γ,n) cross sections reported by the Saskatchewan group[4] from this point of view. With Mg^{25}, Mg^{26}, Si^{29+30}, and A^{40} the (γ,p) resonances were found to have considerably higher energy than (γ,n) resonances. With magnesium the position of the (γ,p) peaks of Mg^{25} and Mg^{26} appear very close to that of the (γ,n) peak of Mg^{24} and the (γ,n) peaks of the former nuclides lie several Mev lower. It is hard to explain these differences by mere Coulomb barrier effect on the protons. It is very tempting to attribute the proton giant resonances to higher T excitation and the neutron resonances to lower T excitation.

A very clear-cut example which might support this view has already been shown by Nathans and Halpern[8] with the case of $Be^9(\gamma,n)$ reaction, where the cross section has two distinct peaks and the higher peak coincides with the (γ,p) peak. The interpretation has been given that the lower peak is due to the odd-neutron excitation and the higher peak is due to the core excitation. The odd-neutron excitation must correspond to $T=\frac{1}{2}$ states; $T=\frac{3}{2}$ states are made only by exciting the core.

The clarity of the situation with Be^9 should be obscured in the case of higher Z, since the binding energy of the odd neutron becomes higher and it becomes more difficult experimentally to find the second peak.

The discussion with Mg^{25} can be extended to the case of Mg^{26}, but in the case of Al^{27} some caution is necessary. From the low proton threshold, it must not be concluded that there is a loosely bound odd proton. Protons are emitted both from $T=\frac{3}{2}$ and $T=\frac{1}{2}$ states, but there is no odd-proton excited state corresponding to odd-neutron excited states in Mg^{25}. Indeed, experiment shows that the (γ,p) peak is as high as 21 Mev even though the (γ,p) threshold is only 7.4 Mev.[7]

The possibility of explaining the separate giant resonances indicates the validity of isotopic spin selection rules in the reaction through highly excited states of light nuclei. The analysis presented here includes higher Z than considered by Gell-Mann and Telegdi.[6] The fact that the (d,α) reaction on Si^{28} with 7-Mev deuterons failed to find the lowest $T=1$, $J=0$ state in Al^{26},[9] also indicated that the isotopic spin selection rules hold in the reaction through 19-Mev states of P^{30}.

Since level spacings are expected to be quite narrow in such high energy states of a nucleus with A as high as 30, it is rather surprising to see the isotopic spin selection rules hold. Probably it is due to the following reasons:

(1) The population of states which have different isotopic spin and which can mix with the state under consideration is much less than the total population of states, since only those states with the same J, L, S, and parity are mixed by Coulomb perturbation.

(2) The perturbation starts as soon as the incident particle or quantum hits the target nucleus. The isotopic spin before the encounter is very well defined since the mixture of higher isotopic spin states in the ground

TABLE I. Measured (γ,p) and (γ,n) cross sections.

	Lithium γ rays[a] (17.6, 14.8 Mev) ratio to $Cu^{63}(\gamma,n)$ yield (reference 1)		Integrated cross section from 26-Mev bremsstrahlung (Mev-barn)[b]		Integrated cross section from 70-Mev bremsstrahlung (Mev-barn)[c]	
	(γ,n)	(γ,p)	(γ,n)	(γ,p)	(γ,n)	(γ,p)
$Mg^{24}(\gamma,n)$	1.1(16.4)[d]		0.057		0.072	
$Mg^{25}(\gamma,p)$		2.83(12.3)		0.10		0.23[e]
$Mg^{26}(\gamma,p)$		1.56(13.0)		0.085		0.092
$Al^{27}(\gamma,n)$[f]	4.0(14.4)		0.045			0.080
$Al^{27}(\gamma,p)$				0.12[g]		
$Si^{29}(\gamma,p)$		3.45(11.7)				
$Si^{30}(\gamma,p)$		1.26(13.7)				
$P^{31}(\gamma,n)$			0.129			

[a] Values in this column should not be taken seriously since the γ rays used are almost monochromatic and the energies are not high enough, so the discussion does not apply so well.
[b] See reference 2.
[c] See reference 3.
[d] The number in parentheses is the threshold value for the reaction.
[e] This value contains a contribution from the $Mg^{26}(\gamma,pn)$ reaction and therefore should be read as about half of this value.
[f] The values cited here are measured by the 6.3-sec activity which is not all of the (γ,n) product. The value must be multiplied by 3. R. Montalbetti and L. Katz, Phys. Rev. 91, 659 (1953).
[g] This value was taken from reference 7.

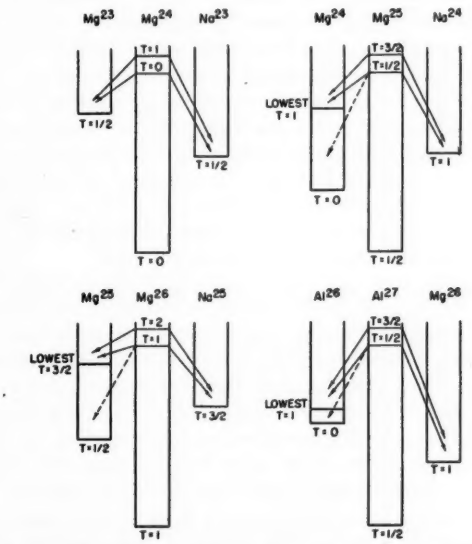


FIG. 1. Schematic diagram showing (γ,p) and (γ,n) reactions from Mg^{24}, Mg^{25}, Mg^{26}, and Al^{27}. The solid arrows show decay to the symmetrical parts of product nuclides. The broken arrows are to the asymmetric part which is reached only by neutron emission.

[8] R. Nathans and J. Halpern, Phys. Rev. 92, 940 (1953).

[9] C. P. Browne, Phys. Rev. 95, 860 (1953).

states is small.[10] A time-dependent perturbation causes the transition of the isotopic spin state, and if H_c is the Coulomb interaction and $\hbar\omega=E_1-E_2$ is the difference in energy between the mixing states, the amplitude of mixing is

$$a(t)=\frac{H_c \sin\omega t}{\hbar\omega}$$

$$\approx\frac{H_c t}{\hbar}=\frac{H_c}{\Gamma}\cdot\frac{t}{T} \text{ (when } t\ll 2\pi/\omega).$$

[10] L. A. Radicati, Proc. Phys. Soc. (London) **A66**, 139 (1953); **A67**, 39 (1954).

Here $\Gamma=h/T$, where T is the life of the decaying state; if this is much larger than H_c, decay takes place before mixing proceeds. Γ is of order 1 Mev or more and H_c is probably a fraction of one Mev. Hence the isotopic spin purity is not affected by close-lying intermixable levels, which would cause considerable mixing in the case of a stationary perturbation.

Further studies on this subject are in progress.

The author is grateful to Dr. D. J. Zaffarano for his constant encouragement. He also wishes to thank Dr. L. Katz for making the latest data available to him, and Drs. D. C. Peaslee and G. Takeda for discussions.

PHYSICAL REVIEW VOLUME 97, NUMBER 2 JANUARY 15, 1955

Energy of the Ground State of Li^6

J. IRVING AND D. S. SCHONLAND
Department of Mathematics, University of Southampton, Southampton, England
(Received September 16, 1954)

The binding energy of Li^6 is calculated by using a wave function of the exponential type with a central Yukawa interaction, both neutral and symmetric exchange characters being considered. The neutral interaction leads to a large excess binding energy whereas the symmetric case gives much too small a value, for a particular set of nuclear parameters. The contribution to the energy from the central part of the neutral and symmetric types of Pease-Feshbach interaction is also determined.

1. INTRODUCTION

IN this paper, the energy of the ground state of Li^6 is determined by using a wave function of the exponential type for a central Yukawa interaction, both the neutral and symmetric cases being considered. The contribution to the energy of the central part of the Pease-Feshbach[1,2] type of interaction—neutral and symmetric—is also evaluated. The justification for dealing only with the central part of the interaction at this stage arises from the treatment of H^3 and He^4, where it is necessary to construct as good a wave function as possible for the S-state,[3] before taking into account the tensor part of the interaction.

It has previously been established that a two-body interaction, involving a mixture of central and tensor forces with a Yukawa well-shape, can give, for a set of nuclear parameters which fits the low-energy two-body data, reasonable values for the binding energies of both the triton[1,2] and the alpha particle.[3] Since the results for the two-body problem are independent of the exchange nature of the forces and the three- and four-body energy values differ very little if a neutral or symmetric interaction is used, it is of importance to determine whether the effect of both neutral and symmetric interactions of the above type is the same for the lightest bound p-shell nucleus, Li^6.

[1] R. L. Pease and H. Feshbach, Phys. Rev. **81**, 142 (1951).
[2] R. L. Pease and H. Feshbach, Phys. Rev. **88**, 945 (1952).
[3] J. Irving, Proc. Phys. Soc. (London) **A66**, 17 (1953).

If the nuclear interaction is assumed to be of the two-body type, saturation requirements for heavy nuclei indicate that the interaction is of a symmetric character. Kronheimer[4] has in fact shown that the exchange nature of the interaction is evident in the case of the light nucleus Be^9. Using single-particle Gauss wave functions and taking only the lowest state (^2P) of highest orbital symmetry of the $(1s)^4(2p)^5$ configuration, he has obtained an excess binding energy with the neutral Pease-Feshbach type of interaction. For the charge-symmetric interaction on the other hand, the $(1s)^4(2p)^5$ term does not describe a bound state. Edwards[5] has found, in the case of the Be^8 nucleus, that for a symmetric central interaction with a Gauss well-shape the system is not bound. Morpurgo,[6] using a similar interaction, has calculated the energy of Li^6, treating the system as composed of a deuteron and an alpha particle. He finds that the energy is a minimum when the deuteron is at infinity, that is, the system is not bound.

Other calculations on the binding energy of Li^6 have been carried out by Inglis,[7] Margenau,[8] and Tyrrell,[9] using a central two-body Gauss interaction, and by

[4] E. H. Kronheimer, Phys. Rev. **90**, 1003 (1953).
[5] S. F. Edwards, Proc. Cambridge Phil. Soc. **48**, 652 (1952).
[6] G. Morpurgo, Nuovo. cimento **10**, 473 (1953).
[7] D. Inglis, Phys. Rev. **51**, 531 (1937).
[8] H. Margenau and K. Carroll, Phys. Rev. **54**, 705 (1938).
[9] W. Tyrrell, Jr., Phys. Rev. **56**, 250 (1939).

Humblet[10] for a Yukawa interaction. Feingold[11] has calculated the level spacing in Li^6 using a central tensor interaction of the Gaussian type. He obtains qualitative agreement with the experimental values for the spacing, but he has not shown that his wave function gives a reasonable value for the binding energy of the Li^6 nucleus. Finally, Morita and Tamura[12] have, independently of the authors, calculated the binding energy of Li^6 with the same exponential wave function for the neutral interaction only, using an elegant adaptation of the method of Jahn and Van Wieringen.[13]

2. METHOD OF CALCULATION

Since Li^6 has spin 1, the ground state on the basis of the shell model is taken to be $^3S_1$, in the usual notation. This state is represented by the wave function, of a similar form to that introduced by Feingold,

$$\begin{aligned}\Psi(^3S_1)=\{&[S(34)\alpha_4A(56)+S(35)\alpha_5A(64)\\&+S(36)\alpha_6A(45)]A(12)\alpha_3+[S(14)\alpha_4A(56)\\&+S(15)\alpha_5A(64)+S(16)\alpha_6A(45)]A(23)\alpha_1\\&+[S(24)\alpha_4A(56)+S(25)\alpha_5A(64)\\&+S(26)\alpha_6A(45)]A(31)\alpha_2\}\phi=\chi\phi,\end{aligned}\tag{2.1}$$

where particles 1, 2, 3 are neutrons and 4, 5, 6 are protons; $S(ij)=\mathbf{r}_{Gi}\cdot\mathbf{r}_{Gj}$, G being the center of mass of the system; $A(ij)=\alpha_i\beta_j-\alpha_j\beta_i$, α and β denoting the usual spin wave functions; ϕ denotes the radial part of the wave function, which is assumed to have the form

$$\phi=N_S^{\frac{1}{2}}\exp\left[-\alpha\left(\sum_{i>j}^{6}{}_1\, r_{ij}^2\right)^{\frac{1}{2}}\right],\tag{2.2}$$

$N_S^{\frac{1}{2}}$ being the normalizing factor for the complete wave function (2.1) and α a variation parameter. The choice (2.2) is suggested by the success of a wave function of this type in accounting for the energy of He^4. Feingold assumes a Gaussian radial dependence.

We consider the interaction of the form

$$\begin{aligned}V^c(r_{ij})=PV_0\{&[(1-\tfrac{1}{2}g)+\tfrac{1}{2}g(\sigma_i\cdot\sigma_j)]\\&\times\exp(-\kappa_c r_{ij})/(\kappa_c r_{ij})\},\end{aligned}\tag{2.3}$$

where $P=-1$ and $g=g_N$ for the neutral case; $P=(\mathbf{T}_i\cdot\mathbf{T}_j)/3$ and $g=g_S=2-3g_N$ for the symmetric case. This interaction is of the same form as the central part of the more general interaction involving the tensor force, considered by Feshbach and Schwinger[14] in the two-body problem and by Pease and Feshbach in the triton problem.

The variational formula,

$$E\leqslant W=\int\Psi^*H\Psi d\tau\Big/\int\Psi^*\Psi d\tau,\tag{2.4}$$

where

$$H=-\frac{\hbar^2}{2M}\sum_{i=1}^{6}\Delta_i+\sum_{i>j}^{6}{}_1\, V^c(r_{ij})$$

and $d\tau$ includes summation over the spins and spatial integrations, is used to obtain an approximate value W to the true energy E. Substituting the wave function (2.1) and the interaction (2.3) in (2.4), the spin matrix elements are first of all evaluated in the usual way. This is tedious but straightforward. The space integrations are then carried out, by means of the transformation of coordinates given by

$$\begin{aligned}\mathbf{x}_1&=(\mathbf{r}_2-\mathbf{r}_1)/\sqrt{2},\\\mathbf{x}_2&=(\mathbf{r}_6-\mathbf{r}_5)/\sqrt{2},\\\mathbf{x}_3&=(\mathbf{r}_6+\mathbf{r}_5-\mathbf{r}_2-\mathbf{r}_1)/2,\\\mathbf{x}_4&=(2/\sqrt{3})\{(\mathbf{r}_3+\mathbf{r}_4)/2-(\mathbf{r}_6+\mathbf{r}_5+\mathbf{r}_2+\mathbf{r}_1)/4\},\\\mathbf{x}_5&=(\mathbf{r}_4-\mathbf{r}_3)/\sqrt{2},\\\mathbf{X}&=(\mathbf{r}_1+\mathbf{r}_2+\mathbf{r}_3+\mathbf{r}_4+\mathbf{r}_5+\mathbf{r}_6)/6,\end{aligned}\tag{2.5}$$

from which

$$\sum_{i>j}^{6}{}_1\, r_{ij}^2=6(x_1^2+x_2^2+x_3^2+x_4^2+x_5^2).\tag{2.6}$$

The method of reducing the resulting spatial integrals and their subsequent evaluation is described in the Appendix. The formulas for W for the neutral and symmetric interactions are now given.

(a) Neutral Interaction ($P=-1$)

$$\begin{aligned}\left\langle\chi^*\left|\sum_{i>j}^{6}{}_1\, V^c(r_{ij})\right|\chi\right\rangle=&-36V_0[\exp(-\kappa_c r_{12})/(\kappa_c r_{12})]\\&\times\{[S^2(12)+6S^2(34)+8S^2(13)-12S(13)S(34)\\&-6S(34)S(35)-6S(13)S(14)+8S(14)S(35)\\&+4S(13)S(24)-2S(14)S(24)-4S(12)S(13)\\&+2S(12)S(34)+S(34)S(56)]+2g_N[-3S^2(34)\\&-2S^2(13)+6S(13)S(34)+3S(34)S(35)\\&+2S(13)S(14)-6S(14)S(35)+S(13)S(24)\\&-S(14)S(24)]\}.\end{aligned}\tag{2.7}$$

Using the transformation (2.5) and integrating over the

[10] J. Humblet, Physica **14**, 285 (1948).
[11] A. M. Feingold, thesis, Princeton University, 1952 (unpublished).
[12] M. Morita and T. Tamura, Tokyo University of Education (private communication).
[13] H. A. Jahn and H. Van Wieringen, Proc. Roy. Soc. (London) A209, 502 (1951).
[14] H. Feshbach and J. Schwinger, Phys. Rev. **84**, 194 (1951).

angles, we then obtain the potential energy term,

$$\left\langle \Psi^* \left| \sum_{i>j}^{6} v^c(r_{ij}) \right| \Psi \right\rangle$$
$$= -[N_S V_0 (4\pi)^5/(\sqrt{2}\kappa_c)]$$
$$\times \int \{[9x_1^4 + 42x_1^2x_2^2 + 267x_2^4/2 - 199x_2^2x_3^2/2]$$
$$- 6g_N[2x_1^2x_2^2 + 21x_2^4 - 17x_2^2x_3^2]\}$$
$$\times \exp\left[-\beta\left(\sum_{i=1}^{5} x_i^2\right)^{\frac{1}{2}} - \sqrt{2}\kappa_c x_1\right] x_1x_2^2x_3^2x_4^2x_5^2$$
$$\times dx_1dx_2dx_3dx_4dx_5 \quad (\text{with } \beta = 2\alpha\sqrt{6})$$
$$= -\frac{5\cdot 11\cdot 13\cdot 17}{9\cdot 2^{14}} V_0 c^{19}\{[369A_1 + 588B_1 + 504C_1]$$
$$- 6g_N[54A_1 + 28B_1]\}, \quad (2.8)$$

on reducing the multiple integrals. Here $c = 2\sqrt{3}\alpha/\kappa_c$ and A_1, B_1, C_1 are single integrals which are given in the Appendix.

The kinetic energy [App. (b)] reduces to

$$3\hbar^2\alpha^2/M = \hbar^2\kappa_c^2c^2/(4M). \quad (2.9)$$

The Coulomb energy [Appendix (c)] becomes

$$(5^2\cdot 11\cdot 13\cdot 17 e^2\kappa_c c)/2^{17}. \quad (2.10)$$

(b) Symmetric Interaction $[P = (\mathbf{T}_i\cdot\mathbf{T}_j)/3]$

$$\left\langle \chi^* \left| \sum_{i>j}^{6} v^c(r_{ij}) \right| \chi \right\rangle$$
$$= -36V_0[\exp(-\kappa_c r_{12})/(\kappa_c r_{12})]$$
$$\times \{[2S^2(34) + S^2(12) - 4S(12)S(13)$$
$$+ 2S(12)S(34) + 2S(14)S(15) + 2S(13)S(23)$$
$$- 4S(13)S(34) - 2S(34)S(35) + S(34)S(56)]$$
$$+ 2g_S[S^2(34) - S(14)S(25) - 2S(13)S(34)$$
$$+ 2S(14)S(35) - S(34)S(35)$$
$$+ S(13)S(23)]\}. \quad (2.11)$$

Hence

$$\left\langle \Psi^* \left| \sum_{i>j}^{6} v^c(r_{ij}) \right| \Psi \right\rangle$$
$$= -\{N_S V_0 (4\pi)^5/(\sqrt{2}\kappa_c)\}$$
$$\times \int \{[9x_1^4 - 30x_1^2x_2^2 + 99x_2^4/2 - 63x_2^2x_3^2/2]$$
$$+ 2g_S[-6x_1^2x_2^2 + 21x_2^4 - 17x_2^2x_3^2]\}$$
$$\times \exp\left[-\beta\left(\sum_{i=1}^{5} x_i^2\right)^{\frac{1}{2}} - \sqrt{2}\kappa_c x_1\right] x_1x_2^2x_3^2x_4^2x_5^2$$
$$\times dx_1dx_2dx_3dx_4dx_5 \quad (\text{with } \beta = 2\alpha\sqrt{6})$$
$$= -\frac{5\cdot 11\cdot 13\cdot 17}{9\cdot 2^{14}} V_0 c^{19}[(153A_1 - 420B_1 + 504C_1)$$
$$+ 6g_S(18A_1 - 28B_1)]. \quad (2.12)$$

The kinetic and Coulomb energy terms are given by (2.9) and (2.10), respectively.

3. RESULTS AND CONCLUSIONS

For the central Yukawa interaction with $1/\kappa_c = 1.17 \times 10^{-13}$ cm, $V_0 = 67.3$ Mev, and $g_N = 0.155$ (i.e., $g_S = 1.535$), we obtain

$W = -101$ Mev and $E_{\text{Coul}} = 4$ Mev, i.e., $E = -97$ Mev

for the neutral case; and

$W = -4.5$ Mev and $E_{\text{Coul}} = 2$ Mev, i.e., $E = -2.5$ Mev

for the symmetric case. E(experimental) $= -32.0$ Mev.

It is evident that the neutral interaction leads to a collapsed nucleus. The very small value for the energy obtained for the symmetric interaction indicates that the assumed form of the wave function is rather poor. The introduction of a tensor force term into the two-body interaction has the effect of reducing the central well depth and consequently the central force binding energy. For the Pease-Feshbach interaction with

$$1/\kappa_c = 1.184\times 10^{-13} \text{ cm}, \quad V_0 = 46.1 \text{ Mev}, \quad g = -0.004,$$
$$\gamma = 0.54, \quad 1/\kappa_t = 1.67\times 10^{-13} \text{ cm},$$

the central force contribution is

$$W = -17.7 \text{ Mev} \quad \text{and} \quad E_{\text{Coul}} = 2.8 \text{ Mev},$$
$$\text{i.e.,} \quad E = -14.9 \text{ Mev},$$

in the neutral case, but the symmetric form of this interaction does not give a bound state.

The tensor force will, of course, contribute to the binding energy. The work of Lyons and Feingold,[15] in which the D state of maximum symmetry, involving a mixture of configurations, is considered, indicates a contribution $\sim$12 Mev to the binding energy. Other D states will give additional binding so that the above neutral central-tensor interaction probably gives excess binding. The symmetric interaction, on the other hand, will still yield little, if any, binding energy for the above wave function. Morita and Tamura claim a fit for the Li^6 ground state energy with a neutral interaction, for which the central depth is 49.3 Mev. Such an interaction will not only give an excess binding energy for the triton and the alpha particle, but will also give an excess for Li^6, if a D state of the same form as that of Feingold is assumed. Morita and Tamura take the D state from the $(1s)^4(2p)^2$ configuration alone and find that it slightly reduces the total energy, in contrast with the result of Feingold and Lyons. Moreover, the work of Cohen[16] on the binding energy of He^4 indicates that the Pease-Feshbach interaction, of a symmetric

[15] D. H. Lyons and A. M. Feingold, Phys. Rev. **95**, 606 (1954).
[16] L. Cohen, thesis, University of Manchester, 1953 (unpublished).

character, with $V_0=42.7$ Mev is the more correct for a reasonable fit in the two-, three- and four-body problems. It should be emphasized that calculations of the Li^6 ground state energy should be carried out with a symmetric interaction.

APPENDIX

(a) Normalization of the Wave Function

The normalization of the wave function involves integrals of the form

$$I=\int_0^\infty\cdots\int_0^\infty f(x_1^2+x_2^2+\cdots+x_n^2)\times x_1^{\alpha_1}x_2^{\alpha_2}\cdots x_n^{\alpha_n}dx_1dx_2\cdots dx_n,$$

which, on using the transformation

$$u_r=x_r^2,\quad r=1,2,\cdots,n, \tag{1}$$

becomes

$$\frac{1}{2^n}\int_0^\infty\cdots\int_0^\infty f(u_1+u_2+\cdots+u_n)\times u_1^{\beta_1}u_2^{\beta_2}\cdots u_n^{\beta_n}du_1du_2\cdots du_n,$$

where

$$\beta_r=(\alpha_r-1)/2,\quad r=1,2,\cdots,n.$$

The further transformation,

$$u_1+u_2+\cdots u_{n-r+1}=y_1y_2\cdots y_r,\quad r=1,2,\cdots,n, \tag{2}$$

yields, with elementary integration,

$$I=\frac{\Gamma(\beta_1+1)\cdots\Gamma(\beta_n+1)}{2^n\Gamma(\beta_1+\cdots+\beta_n+n)}\int_0^\infty y_1^{(\beta_1+\cdots+\beta_n+n-1)}f(y_1)dy_1. \tag{3}$$

For the particular form

$$f(x_1^2+\cdots+x_n^2)=f(y_1)=\exp(-\beta y_1^{\frac{1}{2}}),$$

we have, from (3),

$$I=\pi^{-\frac{1}{2}}\beta^{-n}\left(\frac{2}{\beta}\right)^{\alpha_1+\cdots+\alpha_n}\Gamma\left(\frac{\alpha_1+\cdots+\alpha_n+n}{2}\right)\times\Gamma\left(\frac{\alpha_1+1}{2}\right)\cdots\Gamma\left(\frac{\alpha_n+1}{2}\right). \tag{4}$$

Now for Li^6,

$$\langle\Psi^*|\Psi\rangle$$

$$=36\int[S^2(34)-2S(34)S(35)+S(34)S(15)]\phi^2dv$$

$$=36(4\pi)^5N_S\int_0^\infty\cdots\int_0^\infty\exp[-(2\alpha\sqrt{6})(x_1^2+\cdots+x_5^2)^{\frac{1}{2}}](x_1^4-x_1^2x_2^2)x_1^2dx_1\cdots x_5^2dx_5=1.$$

Hence, using (4) with $n=5$, we obtain

$$N_S=\frac{36\alpha^{19}\sqrt{6}}{35\pi^7}. \tag{5}$$

(b) Kinetic Energy

Using the transformation of coordinates given by (2.5), the kinetic energy is

$$T=(\hbar^2/2M)\sum_{j=1}^5\int\{\nabla_{\mathbf{x}_j}(\chi\phi)\}^2d\tau,$$

where ϕ is the radial, and χ the orbital-spin part of the wave function.

Since $\nabla^2\chi=0$, it follows that

$$T=(\hbar^2/2M)\sum_{j=1}^5\int\chi^2(\nabla_{\mathbf{x}_j}\phi)^2d\tau.$$

On carrying out the spin summations we obtain

$$T=(\hbar^2/2M)36\int\{S^2(34)-2S(34)S(35)+S(34)S(15)\}(\nabla_{\mathbf{x}_j}\phi)^2dv.$$

Now, if ϕ has the form (2.2), then

$$(\nabla_{\mathbf{x}_j}\phi)^2=6\alpha^2\phi^2,$$

i.e.,

$$T=(\hbar^2/2M)36\int\{S^2(34)-2S(34)S(35)+S(34)S(15)\}\phi^2dv\times6\alpha^2.$$

Hence

$$T=3\hbar^2\alpha^2/M \tag{6}$$

since

$$36\int\{S^2(34)-2S(34)S(35)+S(34)S(15)\}\phi^2dv$$

is the normalization integral.

(c) Coulomb Energy

The Coulomb energy is given by

$$E_{\text{Coul}}=3\langle\Psi^*|e^2/r_{56}|\Psi\rangle$$

$$=36e^2\int\{2S^2(35)+S^2(34)-S(35)S(36)-2S(34)S(35)-2S(35)S(15)-S(34)S(14)+S(35)S(16)+2S(35)S(14)\}(\phi^2/r_{56})dv,$$

which gives, when one uses the preceding integrals,

$$E_{\text{Coul}}=(5^2\cdot11\cdot13\cdot17e^2c\kappa_c)/2^{17}. \tag{7}$$

(d) Potential Energy

The potential energy terms involve integrals of the form

$$J=\int_0^\infty\cdots\int_0^\infty f(x_1^2+\cdots+x_n^2)g(x_n)\times x_1^{\alpha_1}\cdots x_n^{\alpha_n}dx_1\cdots dx_n. \quad (8)$$

When one uses the transformations (1) and (2), Eq. (8) reduces to

$$\frac{\Gamma(\beta_1+1)\Gamma(\beta_2+1)\cdots\Gamma(\beta_{n-1}+1)}{2^n\Gamma(\beta_1+\cdots+\beta_{n-1}+n-1)}\times\int_0^\infty dy_1 f(y_1)y_1^{(\beta_1+\beta_2+\cdots+\beta_n+n-1)}\times\int_0^1 dy_2 g\{y_1^{\frac{1}{2}}(1-y_2)^{\frac{1}{2}}\}(1-y_2)^{\beta_n}, \quad (9)$$

where

$$\beta_r=(\alpha_r-1)/2.$$

For the particular forms,

$$f(y_1)=\exp(-\beta y_1^{\frac{1}{2}}),$$

$$g\{y_1^{\frac{1}{2}}(1-y_2)^{\frac{1}{2}}\}=\exp\{-\kappa' y_1^{\frac{1}{2}}(1-y_2)^{\frac{1}{2}}\}/\{\kappa' y_1^{\frac{1}{2}}(1-y_2)^{\frac{1}{2}}\},$$

we obtain from (9), on carrying out the integration with respect to y_1,

$$J=\frac{\Gamma(\beta_1+1)\cdots\Gamma(\beta_{n-1}+1)\Gamma\{2(\beta_1+\cdots+\beta_n+n)-1\}}{\kappa'^{2n-1}\Gamma(\beta_1+\cdots+\beta_{n-1}+n-1)}\times\int_0^1 dy_2\frac{y_2^{\beta_1+\cdots+\beta_{n-1}+n-2}(1-y_2)^{\beta_n-\frac{1}{2}}}{[\beta+\kappa'(1-y_2)^{\frac{1}{2}}]^{[2(\beta_1+\cdots+\beta_n+n)-1]}}. \quad (10)$$

Using (10), with $n=5$, the potential energy is determined and involves the integrals A_1, B_1, C_1 where

$$A_1=\int_0^1(1-y^2)^7y/(y+c)^{18}dy, \quad (11)$$

$$B_1=\int_0^1(1-y^2)^6y^3/(y+c)^{18}dy, \quad (12)$$

$$C_1=\int_0^1(1-y^2)^5y^5/(y+c)^{18}dy, \quad (13)$$

with $c=2\sqrt{3}\alpha/\kappa_c$. Evaluating A_1, B_1, C_1 by elementary methods, we obtain

$$(5\cdot9\cdot11\cdot13\cdot17\cdot2^4)c^{16}(1+c)^{10}A_1=[109\,395c^8+504\,126c^7+1\,103\,040c^6+1\,472\,130c^5+1\,293\,930c^4+759\,330c^3+288\,288c^2+64\,350c+6435],$$

$$(3\cdot5\cdot7\cdot11\cdot13\cdot17\cdot2^4)c^{14}(1+c)^{11}B_1=[34\,465c^7+106\,203c^6+147\,213c^5+123\,735c^4+67\,155c^3+23\,265c^2+4719c+429],$$

$$(7\cdot9\cdot11\cdot13\cdot17\cdot2^4)c^{12}(1+c)^{12}C_1=[7293c^6+13\,788c^5+12\,303c^4+6504c^3+2115c^2+396c+33].$$

ACKNOWLEDGMENT

The authors wish to thank M. Morita and T. Tamura for their courtesy in sending them a prepublication copy of their paper.

Radioactive Decay of Cs^{134} and Cs^{134m}†

G. L. KEISTER,* E. B. LEE,‡ AND F. H. SCHMIDT
Department of Physics, University of Washington, Seattle, Washington
(Received July 26, 1954)

The radiations of Cs^{134} (2.3 yr) and Cs^{134m} (3.15 hr) have been studied in a high-resolution beta-ray spectrometer. The gamma rays of Cs^{134} were observed both by internal and external conversion, and coincidence rates were measured between the beta continuum and the internal conversion electrons of the stronger gamma rays. The beta spectrum of Cs^{134} appears to consist of four, or possibly five, components. Nine gamma rays were found both by internal and external conversion; a tenth gamma ray appears only in internal conversion. A decay scheme is proposed which is reasonably consistent with the multipole order of each of the radiations as obtained from internal conversion data.

A search was made for a ground- or intermediate-state beta transition from Cs^{134m}. No transition to the ground state was observed, but some indication was found for a weak transition to an excited state.

I. INTRODUCTION

THE first detailed study of the radiations from 2.3-year Cs^{134} were made by Elliot and Bell.[1] Since then several additional researches have appeared in the literature,[2] each of which has indicated further complexity in the decay scheme. Waggoner, Moon, and Roberts[3] added three weak high-energy gamma rays to the three strong ones found previously. Schmidt and Keister[4] found that one of the gamma rays was actually two closely-spaced rays, and accordingly proposed a new decay scheme.[5] More recently Cork *et al.*[6] reported a total of eleven gamma rays resulting from the decay of Cs^{134}. These workers proposed still another decay scheme.

We have made a further study of Cs^{134} in an effort to resolve the several inconsistencies in the complicated decay scheme. In this objective we have only been partially successful, for we have yet to find a scheme which is wholly consistent with the experimental measurements. The radiations were examined in a high-resolution ring-focus beta-ray spectrometer.[7] The relative intensities and the conversion coefficients of most of the gamma rays have been determined, the beta spectrum has been resolved into several groups, and coincidences between beta particles and internal conversion electrons have been studied in the spectrometer. These data, together with angular correlation and related measurements,[2] suggest a modified decay scheme for which reasonably consistent spins and parities can be assigned to the various energy levels.

II. THE GAMMA RAYS AND CONVERSION COEFFICIENTS

The relative intensity of each gamma ray was determined by means of photoelectric conversion in thin thorium foil radiators. The mechanical arrangement of the radioactive material and radiator is shown in Fig. 1. A small Dural capsule containing the active material is made removable, and therefore the same foil and identical geometry can be used with various sources of gamma radiation.

The relative intensity of a photoconversion line in the spectrometer when there is a space distribution of source material and an extended radiator foil will, in general, depend upon the gamma-ray energy in a direct manner, and also in an indirect manner through energy-dependent angle factors.[8] For our particular source and radiator geometry, the dependence of the intensity upon energy through the angle factors is expected to be small in the range 0.5 to 1.5 Mev. In order to verify this conclusion an experimental determination of the

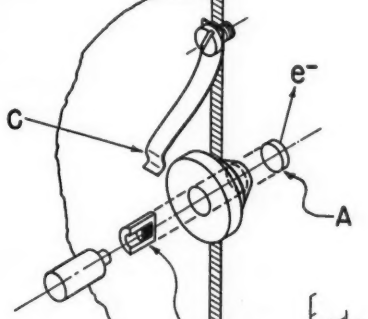


FIG. 1. Photoconversion electron source. (A) Thorium foil radiator; (B) Capsule containing radioactivity; (C) Spring clip to hold source capsule in position.

† Supported in part by the U. S. Atomic Energy Commission.
* Now at Boeing Airplane Company, Seattle, Washington.
‡ Now at Bell Telephone Laboratories, New York, New York.
[1] L. G. Elliot and R. E. Bell, Phys. Rev. **72**, 979 (1947).
[2] K. Way, *Nuclear Data*, National Bureau of Standards Circular No. 499 (U. S. Government Printing Office, Washington, D. C., 1950).
[3] Waggoner, Moon, and Roberts, Phys. Rev. **80**, 420 (1950).
[4] F. H. Schmidt and G. L. Keister, Phys. Rev. **86**, 632 (1952).
[5] Hollander, Perlman, and Seaborg, Revs. Modern Phys. **25**, 469 (1953).
[6] Cork, LeBlanc, Nester, Martin, and Price, Phys. Rev. **90**, 444 (1953).
[7] F. H. Schmidt, Rev. Sci. Instr. **23**, 361 (1952). For recent modifications of the instrument, see reference 10.
[8] W. Heitler, *The Quantum Theory of Radiation* (Oxford University Press, London, 1944).

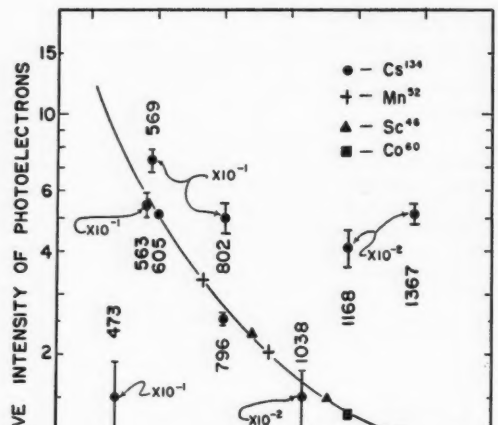


FIG. 2. Gamma-ray intensities of Cs^{134} by photoelectric conversion in thorium foil radiators. The curve is the energy-dependent portion of the theoretical differential cross section.

relative line intensity *vs* energy was made for our particular geometry in the following manner: Sources were studied for which two or more gamma rays are known to be in simple cascade. These include[2] Mn^{52}, Sc^{46}, and Co^{60}, which emit gamma rays of energies that conveniently overlap. The results, as shown in Fig. 2, indicate that the calculated intensity (assuming only the direct energy dependence) is in satisfactory agreement with the measured values over a range from ~0.5 to 1.5 Mev.

High specific activity Cs^{134}, Co^{60}, and Sc^{46} were obtained from Oak Ridge. The Mn^{52} was made in the University of Washington sixty-inch cyclotron.[9]

Of the eleven gamma rays in Cs^{134} reported by Cork *et al.*[6] we were able to resolve and to measure the relative intensities of nine by photoconversion. See Table I. The *K*-shell line of the 663-kev gamma ray, if present, would be partially obscured by the *L*-shell line due to the 569-kev gamma ray. We were unable to observe the very weak line at 202 kev reported by Cork *et al.*[6] either by external or internal conversion. Typical external conversion spectra are shown in Fig. 3. The relative intensity of each gamma ray, normalized to the 605-kev line, is plotted in Fig. 2.

For study of the beta and internal conversion spectra, thin sources of $CsNO_3$ (specific activity >3 curies/g) were prepared by evaporation from a quartz crucible *in vacuo* onto rubber hydrochloride films of surface density ~0.6 mg/cm². The sources ranged from 16 to 40 micrograms/cm² in average surface density. The uniformity and thinness of the source material is indicated by the absence of any measurable "break" at 600 kev due to Compton electrons produced in the source by the strong 794-kev gamma ray.[10] See Fig. 6.

Several internal conversion spectrum lines are shown in Fig. 4. The spectrometer transmission and luminosity were increased for observation of the weaker gamma rays. Accordingly, the resolution is not the same for all lines. For the weakest lines the full line width at half-maximum was ~2.4 percent.

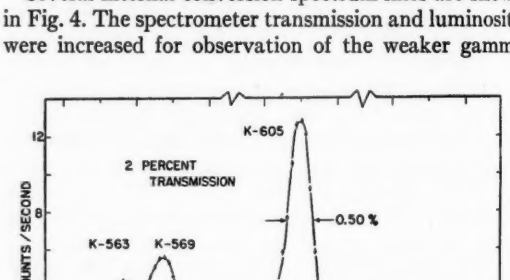


FIG. 4. Representative internal conversion spectrum of Cs^{134}. These lines occur near the end of the continuum. Note that the distance between the lines has been contracted on the graph.

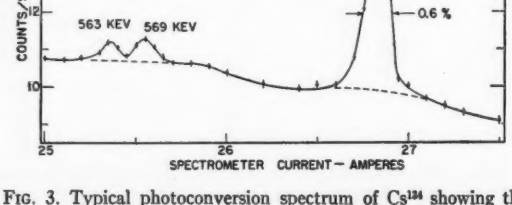


FIG. 3. Typical photoconversion spectrum of Cs^{134} showing the *K*-conversion lines of three of the gamma rays.

The results of the internal and external conversion measurements are summarized in Table I. Our energy values,[11] as tabulated in column 2, are in most cases weighted averages obtained from *K*-shell photoconversion in thorium, *K*-shell internal conversion in barium, and *L*-shell internal conversion in barium. The values published by Cork *et al.*[6] are listed in column 1. Relative intensities of the gamma rays and of the internal conversion lines appear in columns 3 to 6.

Although the evidence is not quite conclusive, both the *K-L* energy differences and the photoelectric conversion data for the three high-energy weak gamma rays

[9] Schmidt, Farwell, Henderson, Morgan, and Streib, Rev. Sci. Instr. **25**, 499 (1954).

[10] G. L. Keister and F. H. Schmidt, Phys. Rev. **93**, 140 (1954). Figure 5 of this paper shows the effect of Compton electrons very clearly.

[11] These differ from the ones published earlier in reference 4. The discrepancy is apparently due to a concurrent new determination of the energy of the *F* line of thorium B which was used as a standard.

support the hypothesis that these gamma rays arise from transitions in barium. This result is in agreement with Cork *et al.*[6] K capture is further excluded by the absence of any measurable Auger electrons. We conclude that a K capture branch cannot exceed one percent. In addition, Mims and Halban[12] placed an upper limit of 0.009 percent on positron emission.

K-shell internal conversion coefficients for each gamma ray are plotted in Fig. 5, and listed in column 2, Table II. These values were obtained by assuming that the 605-kev transition leads to the ground state, and that this radiation is pure $E2$. The coefficient for the 605-kev gamma ray is thus normalized to the theoretical value.[13] We justify the assumption on the following grounds: (1) Angular correlation measurements[14] strongly support a 4-2-0 spin assignment for the levels giving rise to the strongest gamma rays (605 and 796 kev); (2) The coefficient for the 796-kev gamma ray so obtained is thus consistent with an $E2$ assignment; (3) The assignments for other gamma rays is

TABLE I. Cs^{134} gamma rays and conversion lines.

Gamma ray energy-kev		Relative intensity	Relative intensity of internal conversion electrons		
Cork *et al.*[a]	Present		K	L	M
202.5	···	<0.2	<1.2		
475.0	473±2	1.8±0.5	3±1		
563.0	563±1	9.4±2.0	11.9±0.6	—[b]	
569.7	569±1	12.8±2.0	26±1	—[b]	
605.4	605±1	100	100[c]	13±1	2.6±0.6
662.7	663±2[d]	···	1.0±0.5		
796.8	796±1	91±4	44±1	6.3±0.6	2.1±0.4
802.6	801±2	18±4	3.9±0.5		
1039	1038±5	0.9±0.2	0.43±0.08	0.06±0.02[e]	
1168	1168±5	3.0±0.4	0.49±0.08	0.08±0.02[e]	
1368	1367±5	4.6±0.3	0.49±0.08	0.05±0.02[e]	

[a] See reference 6.
[b] Cannot be resolved from the K-line of the 605-kev gamma ray.
[c] Corrected for an assumed K/L ratio of ~6 for both the 563- and the 569-kev gamma rays.
[d] Assuming K-conversion in barium.
[e] $L+M$ conversion.

then unambiguous, as can be noted in Fig. 5; (4) The measured K/L ratios are consistent with the extrapolated empirical[15] and theoretical[16] values for $E2$ radiation for both the 605- and the 796-kev gamma rays. The relative gamma-ray intensities and the coincidence measurements (see Sec. IV) support the contention that the 605-kev transition leads to the ground state.

The K-conversion coefficient for the 605-kev gamma ray has also been calculated by assuming that all beta decays lead to this transition. The result, as shown in Fig. 5, gives a coefficient which is ~20 percent less than the theoretical $E2$ value. This is taken as evidence that another gamma ray leads to the ground state in about

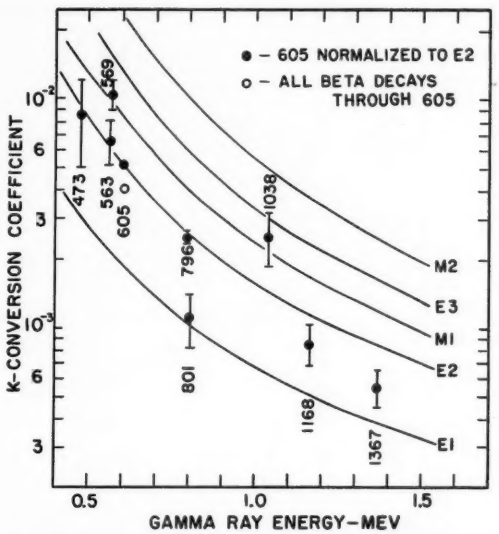


FIG. 5. The K-shell internal conversion coefficients of nine gamma rays in the decay of Cs^{134}. The curves are the theoretical values as obtained from reference 13. The solid points are normalized to the theoretical value for the 605-kev gamma ray. The open circle is the value obtained for the 605-kev gamma if one assumes that all beta decays lead through it.

one out of 5 to 8 disintegrations. The measured intensity of the 801-kev gamma ray indicates that it can play this role.

The other K internal conversion coefficients listed in Table II are plotted in Fig. 5. The solid lines of Fig. 5 are the theoretical[13] values for the various types of radiation. The most probable assignment for each gamma ray, as obtained from Fig. 5, is listed in Table II, column 5.

III. BETA-RAY SPECTRA

The beta-ray continua of both strong and weak Cs^{134} sources (Sec. II) were studied in the spectrometer. The region of the spectrum between the conversion lines which lie very close to the end point was examined at ~1 percent inverse momentum resolution. Detailed measurements of the last 100 kev of the spectrum were

TABLE II. Conversion coefficients of Cs^{134} gamma rays.

Gamma ray energy-kev	K coefficient[a] $\times 10^3$	Experimental K/L	Theoretical[b] K/L_{I+II}	Multipole assignment
473	8.6 ±3.6			$E2$
563	6.6 ±1.5			$E2$
569	10.5 ±1.6			$M1$
605	5.19	7.7±0.6	8.4	$E2$
796	2.51±0.15	8.0±0.8	7.9	$E2$
801	1.1 ±0.3			$E1$
1038	2.5 ±0.7	~ 7[c]		$E3$ or $M1$
1168	0.85±0.18	~ 6[c]		$E1$ or $E2$
1367	0.55±0.10	~10[c]		$E1$ or $E2$

[a] Normalized to theoretical $E2$ value for 605-kev transition.
[b] For $E2$ radiation. See reference 16.
[c] $K/(L+M)$.

[12] W. Mims and H. Halban, Proc. Phys. Soc. (London) **A64**, 311 (1951).
[13] Rose, Goertzel, and Perry, Oak Ridge National Laboratory Report 1023, 1951 (unpublished).
[14] B. L. Robinson and L. Madansky, Phys. Rev. **84**, 604 (1951).
[15] M. Goldhaber and A. W. Sunyar, Phys. Rev. **83**, 906 (1951).
[16] Rose, Goertzel, and Swift, Oak Ridge National Laboratory Report (unpublished). We are indebted to Dr. Rose for kindly supplying us with these data in advance of the full report.

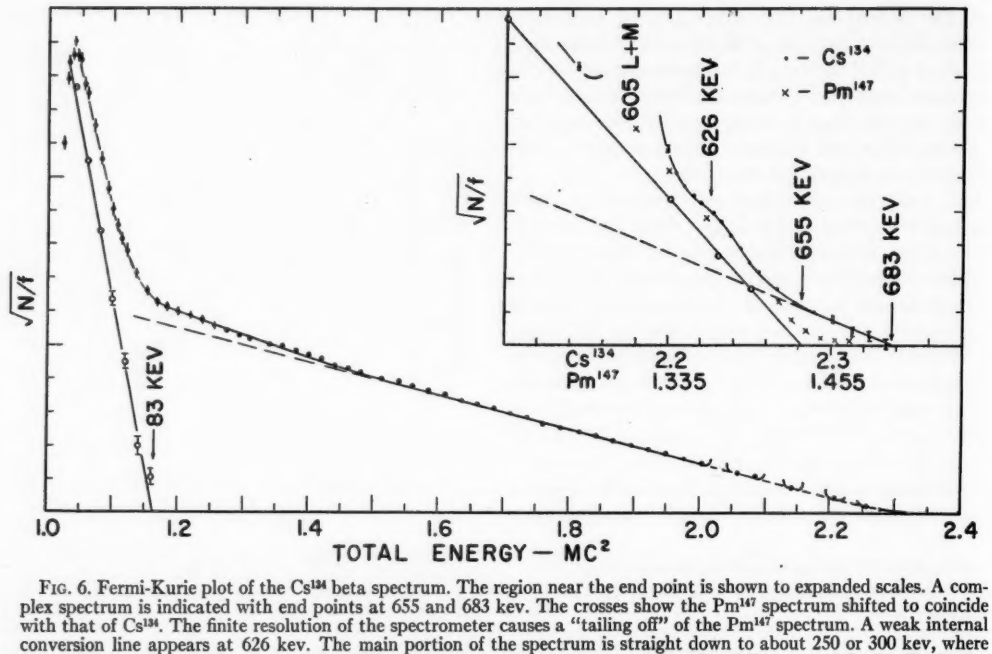


FIG. 6. Fermi-Kurie plot of the Cs^{134} beta spectrum. The region near the end point is shown to expanded scales. A complex spectrum is indicated with end points at 655 and 683 kev. The crosses show the Pm^{147} spectrum shifted to coincide with that of Cs^{134}. The finite resolution of the spectrometer causes a "tailing off" of the Pm^{147} spectrum. A weak internal conversion line appears at 626 kev. The main portion of the spectrum is straight down to about 250 or 300 kev, where one or two weak components may begin. The 83-kev component is clearly evident. It exhibits a straight line Kurie plot down to ~20 kev.

made with a strong source and with the spectrometer baffles set for 4 percent transmission giving an inverse resolution of ~1.7 percent. A Fermi-Kurie plot of the results is shown in Fig. 6. The well-known principal components have end-point energies of 655 and 83 kev. The intermediate region of the spectrum may indicate the existence of two weak components as suggested by Cork *et al.*[6] We have resolved the spectrum into only one additional component in this region.

The inset of Fig. 6 shows the region of the end point of the Cs^{134} spectrum plotted to expanded scales. For comparison purposes the end point of the Pm^{147} spectrum is also shown. At 626 kev there appears a weak conversion line which, if converted in barium, corresponds to a gamma ray of 663 kev. Just to the right of this line there is a definite "break" which may indicate the existence of another component spectrum. There exists the possibility that the *L*-conversion line of the weak 663-kev gamma ray is responsible for the apparent break in the continuum. If this indeed be the case, then the intensity of the *L* line is about equal to the intensity of the *K* line. For a gamma ray of this energy a K/L ratio of one is very improbable.[15] It is tempting to associate the 663-kev gamma ray with the 662-kev gamma ray of Cs^{137} which has a K/L conversion ratio[5] of ~4. The effect on the beta continuum due to a Cs^{137} contaminant of amount thus indicated would be negligible. No changes have occurred in the details of the spectrum over a period of four months which could not be accounted for by the 2.3-year half-life of Cs^{134}.

The end-point energies, relative intensities and log*ft* values of the component spectra are tabulated in Table III.

TABLE III. Cs^{134} beta spectra.

End-point energy in kev		Percent intensity		Log*ft*
Cork *et al.*[a]	Present	Cork *et al.*[a]	Present	(present)
	683±4		13±6	9.9
657±5	655±2	81	50	9.2
410±5	308±100	6	5±2	9.0
210		3		
80	83±3	10	32±6	6.4

[a] See reference 6.

IV. COINCIDENCE MEASUREMENTS

Coincidences between the beta particle continuum and internal conversion electrons were studied in the beta-ray spectrometer with the following apparatus: A trans-stilbene crystal, approximately 0.5 mm thick, was placed about one-half inch behind the beta-ray source. The light from scintillations produced in the crystal were conducted through a 40-inch long Lucite rod to the face of an RCA 5819 photomultiplier tube. The latter was located in a suitable magnetic shield outside the spectrometer vacuum. Only beta particles of energies greater than 90 kev were detected. The lowest energy beta-ray component was therefore not counted.

Another scintillation crystal about $\frac{1}{8}$ inch thick and associated "light piper" was provided at the focus of

the spectrometer in place of the customary G-M tube detector. Coincidence pulses between the two photomultipliers were detected by a circuit of resolving time ~0.2 μsec.

The coincidence measurements show that a larger fraction of the 796- than the 605-kev gamma rays follow high-energy beta decays. The measured ratio of 796 to 605 coincidence rates is 1.3±0.2. The measurements show that very few, if any, of either the 563- or the 569-kev gamma rays are in coincidence with high-energy beta particles.

A search was made with NaI(Tl) scintillation counters for delayed gamma-gamma coincidences. None was observed with a half-life greater than 0.2 μsec. However, our measurements do not quite exclude the possibility of a delay in one of the weaker transitions.

V. Cs^{134m}

The spectrum of the 3.15-hour isomer of Cs^{134} was studied in the spectrometer to determine whether beta transitions occur from the isomeric state to one of the excited states in Ba^{134}. Cs^{134m} was made by the $Cs^{133}(d,p)$ reaction in the University of Washington cyclotron.[9] The strength of the beta-ray source was estimated from the known value[17] of the K-conversion coefficient and the measured intensity of the K-conversion line of the 128-kev transition. A very weak continuous spectrum was observed with an end-point energy ~550 kev, and with a relative intensity ~1 percent. No transition to the ground state of Ba^{134} was found.

VI. DISCUSSION OF DECAY SCHEMES

Figures 7 and 8 show, respectively, the decay scheme proposed by Cork *et al.*,[6] and our modified scheme which is somewhat more consistent with our data.

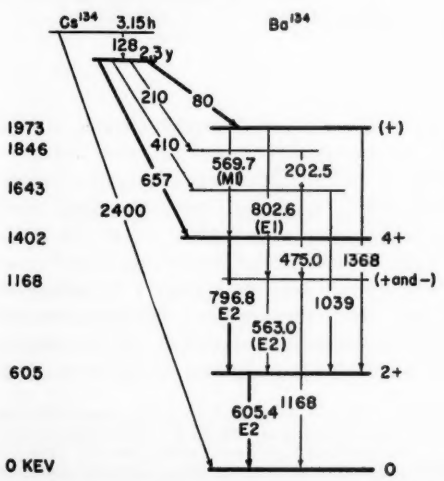


FIG. 7. Decay scheme proposed by Cork *et al.*, reference 6. The multipolarity and parity assignments enclosed in parentheses have been added by us. On the basis of these assignments the level at 1168 kev must be both even and odd.

[17] A. W. Sunyar, Phys. Rev. 83, 864 (1951).

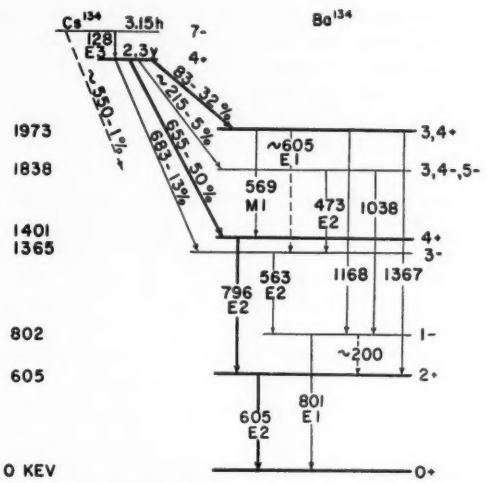


FIG. 8. Newly proposed decay scheme. In order to explain our coincidence and internal conversion measurements, we postulate the existence of a gamma ray of approximately 605-kev energy (shown dotted). The decay scheme for the 3.15-hour isomeric state has recently been shown to consist of a two-step transition [Sunyar, Mihelich, and Goldhaber, Phys. Rev. 95, 570 (1954)].

The scheme of Fig. 7 is consistent with the observed coincidences between beta particles and internal conversion electrons of the 605- and the 796-kev gamma rays. Also note that, according to this scheme, neither the 563- nor the 569-kev gamma rays are in coincidence with high-energy beta particles. However, about 97 percent of all decays should give rise to the 605-kev gamma ray. The internal conversion coefficient so obtained should then be consistent with $E2$ radiation; as mentioned in Sec. II this is not the case. Further, multipolarity assignments made on the basis of the internal conversion coefficients lead to an inconsistent parity assignment for the second excited state.

The decay scheme shown in Fig. 8 is somewhat more consistent with experimental data than that of Fig. 7. However, it contains a distasteful artifice: In order to understand both the coincidence data and the relative gamma-ray intensities, it seems necessary to postulate the existence of a gamma transition between the 1973- and the 1365-kev levels; its energy would necessarily lie within 2 or 3 kev of the 605-kev gamma, and the conversion coefficient would have to be smaller than that for $E2$ radiation. Further support for this postulate is found by comparing the relative intensities of the 83-kev beta transition with that of the 569-kev gamma ray. The "new" 605-kev gamma ray would thus arise from an $E1$ transition of relative intensity 5 to 15 percent. The existence of the "new" 605-kev gamma would depress the value of the conversion coefficients for the more intense 605-kev gamma ray by less than 10 percent. The still lower value K-conversion coefficient obtained for the 605 when all beta transitions are assumed to pass through the 605 is explained by paralleling the 801 with the 605.

The proposed scheme of Fig. 8 with the newly postulated ~605-kev transition encounters difficulties when one considers the deexcitation of the 1365-kev level by only the 563-kev gamma ray. The situation would be partially alleviated if the 1367-kev gamma ray led from the 1365 level to the ground state. The measured conversion coefficient is then too small to be consistent with the spin assignments. In addition, the coincidence measurements indicate that more low-energy beta decays are followed by 605-kev gammas than by 796-kev gammas. The 1367-kev transition as shown has the appropriate "shunting" effect. For this reason too it is difficult to justify inverting the order of the 796 and 605 gammas, although such an inversion would improve the consistency of the 1038 and 1168 gamma rays.

A ground-state spin of 4 has been measured for Cs^{134} by Bellamy and Smith.[18] An even parity is predicted by the shell model. The principal beta transitions shown in Fig. 8 are therefore allowed and the high logft values must be assumed to result from l-forbiddenness. Similar conclusions apply to the two weaker transitions.

The newly-proposed decay scheme is consistent with the lack of coincidences between high-energy beta particles and the 563-kev internal conversion electrons only if the 683-kev beta component is assumed to be less than ~10 percent of all beta transitions.

We wish to thank T. J. Morgan and the cyclotron crew for the production of Mn^{52} and Cs^{134m}, and J. R. Penning for assistance with the delayed gamma-ray measurements.

[18] E. H. Bellamy and K. F. Smith, Phil. Mag. (7) 44, 33 (1953).

PHYSICAL REVIEW VOLUME 97, NUMBER 2 JANUARY 15, 1955

Deuterium He^3 Reaction*

W. E. Kunz†
Oak Ridge National Laboratory, Oak Ridge, Tennessee
(Received September 29, 1954)

By the use of accelerated He^3 ions, the reaction $H^2(He^3,p)He^4$ has been studied in the energy range 100–800 kev (He^3 energy). The angular distribution of the protons was found to be isotropic at bombarding energies of 200, 290, and 350 kev. The reaction cross section has a peak of 695±14 millibarns as determined with reference to the peak value 5.00 barns for the comparison reaction, $H^2(H^2,n)He^4$. The comparison was effected by alternately bombarding the same deuterium target with tritium and He^3 ions, and counting the alpha particles. The peak occurred at 640-kev He^3 ion bombarding energy.

From considerations of the absolute value of the cross section, it is concluded that the resonance at 640 kev is associated with a $J=\frac{3}{2}$ level in the compound nucleus. The experimental shape of the peak is well fitted by resonance parameters in the one-level dispersion formula having the following values: interaction radius R, 5×10^{-13} cm; reduced width for proton emission γ_p, 41.9 kev; reduced scattering width γ_α, 2930 kev; "formal" resonance energy in the center-of-mass system ϵ_s, 391 kev; energy of the level above the ground state of Li^5: 16.2±0.3 Mev. These values agree within experimental error with analogous values associated with the similar resonance which occurs in the $H^2(H^3,n)He^4$ reaction.

INTRODUCTION

THE reaction $H^2(He^3,p)He^4$, which we will speak of simply as "the He^3 reaction" is in a sense the mirror image of the $H^2(H^3,n)He^4$ reaction, which we will herein call "the tritium reaction." The former involves Li^5 (three protons and two neutrons) as the intermediate nucleus, and the latter He^5 (two protons and three neutrons). If nuclear forces are charge symmetric, then the two intermediate nuclei should possess energy levels equivalent in character, differing in fact only in Coulomb energy. It is therefore of interest to compare the cross sections of the two reactions as accurately as possible to see if the known resonances have comparable characteristics. It is also of interest to compare the shapes of the resonance peaks with the shapes predicted by the one-level dispersion formula to see if this kind of theoretical treatment can have validity when applied to nuclei possessing as few as five nucleons.

The tritium reaction has been extensively studied in the energy region below 1 Mev,[1–6] with the result that the peak at 163 kev has been characterized in a satisfactory manner both with respect to reduced width and to absolute cross section. The analogous peak in the He^3 reaction is, however, perhaps less well known, although it also has been the subject of a number of

* Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at the University of Tennessee.

† Oak Ridge Institute of Nuclear Studies Fellow now at the Naval Research Laboratory, Washington, D. C.

[1] Baker, Holloway, King, and Schreiber, Atomic Energy Commission Declassified Report AECD No. 2226, 1948 (unpublished).
[2] E. Bretscher and A. P. French, Phys. Rev. 75, 1154 (1949).
[3] D. L. Allan and M. J. Poole, Proc. Roy. Soc. (London) A204, 488–500 (1951).
[4] Conner, Bonner, and Smith, Phys. Rev. 88, 468 (1952).
[5] Stovall, Arnold, Phillips, Sawyer, and Tuck, Phys. Rev. 88, 159(A) (1952).
[6] Argo, Taschek, Agnew, Hemmendinger, and Leland, Phys. Rev. 87, 612 (1952).

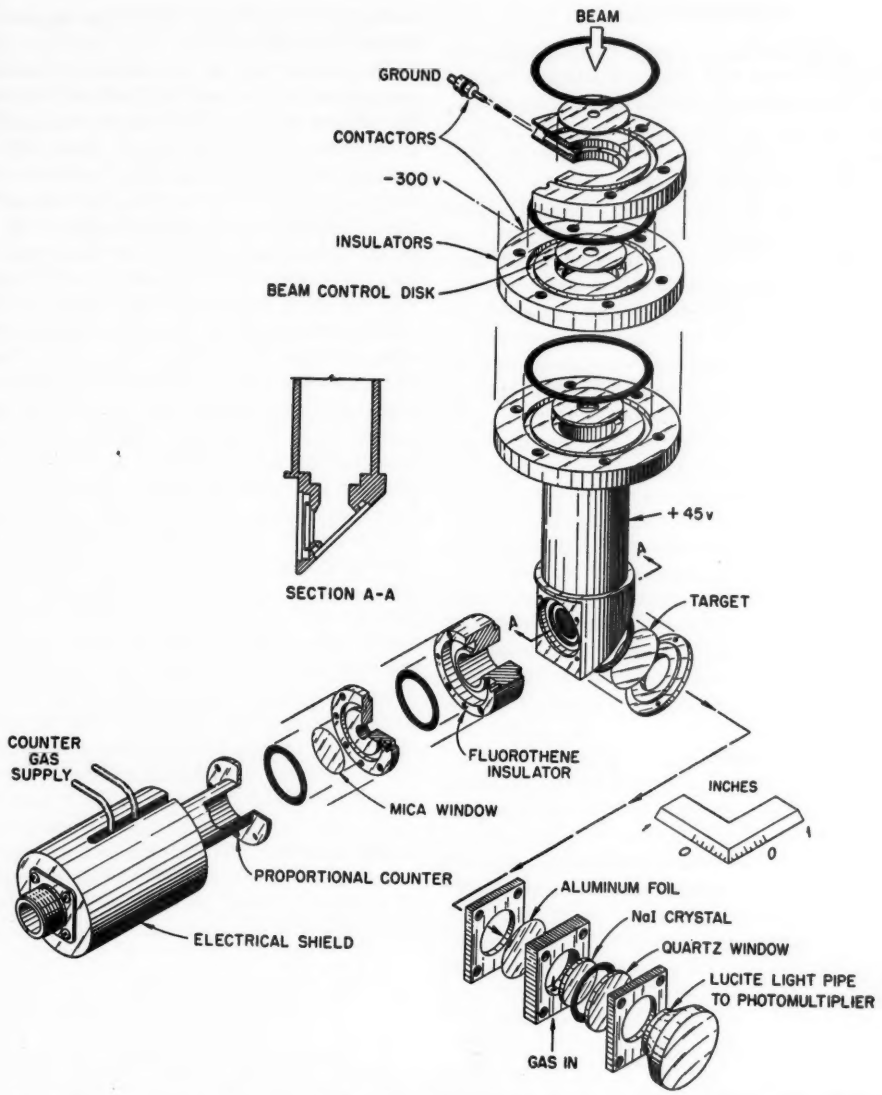


FIG. 1. The target apparatus and counters. Alphas from both the He^3 reaction and the tritium reaction were counted by the proportional counter. The NaI crystal-photomultiplier combination was used to count protons from the He^3 reaction.

investigations.[7–13] In much of the experimental work of this kind, a deuterium beam has been used, with gas targets. In the experiments to be described here, the reversed procedure was used; a common solid deuterium target was bombarded in turn with tritium and with He^3 ions. Accepting the absolute value of the cross section for the tritium reaction as measured by others, it was thus possible to obtain the He^3 reaction cross section by a comparison which kept the target and detecting equipment unchanged while switching from one reaction to the other. With the new information on the He^3 reaction thus obtained, it has been possible to make significant comparisons of the parameters associated with the analogous levels in He^5 and Li^5, and to examine the shapes of the resonance peaks in the light of the dispersion theory.

[7] Baker, Holloway, King, and Schreiber, Atomic Energy Commission Declassified Report AECD No. 2189 (1948).
[8] J. Hatton and G. Preston, Nature 164, 143 (1949).
[9] Tuck, Arnold, Phillips, Sawyer, and Stovall, Phys. Rev. 88, 159(A) (1952).
[10] Bonner, Conner, and Lillie, Phys. Rev. 88, 473 (1952).
[11] Yarnell, Lovberg, and Stratton, Phys. Rev. 90, 292 (1953).
[12] R. G. Jarvis and D. Roaf, Proc. Roy. Soc. (London) A66, 310 (1953).
[13] G. Freier and H. Holmgren, Phys. Rev. 93, 825 (1954).

EXPERIMENTAL

The He^3 and H^3 ion beams used in these experiments were produced in the Oak Ridge Cockcroft-Walton accelerator. The accelerator yields a potential of 400 kv, but it was possible to take advantage of a weak,[14] doubly-charged component of the He^3 beam to carry the bombarding energies to 800 kev.

The first part of the experiment consisted of the determination, on a nonabsolute basis, of the shape of the yield curve of the He^3 reaction from 100 to 800 kev. Figure 1 shows the arrangement of the target and detecting apparatus. The latter consisted of a proportional counter, set at 90 degrees to the incident beam, which detected the alpha particles but was biased against registering the protons, and in addition a NaI crystal-5819 photomultiplier combination, at 45 degrees, used to count the protons only. The NaI crystal subtended 20 times the solid angle presented by the counter to the source, so that it was useful in regions of low yield. Elsewhere the two detectors could be used simultaneously. The number of bombarding He^3 particles was determined by current integration, a 300-volt bias (Fig. 1) preventing the escape of secondary electrons. Charge exchange in the incident ion beam was reduced by keeping the residual gas occupying the 75-cm distance between the beam-bending magnet of the accelerator and the target at a pressure below 5 or 6×10^{-6} mm of mercury.

A thin target was needed in the yield curve determination because the energy loss rate of He^3 ions of these energies in the target material is not well known. Deuterated polyethylene of thickness about $1\mu g/cm^2$ was used; this would present a stopping equivalence of about 1 kev to 300 kev He^3 ions, so the uncertainties involved would be small and the effective He^3 ion energy would be well determined. The targets had the disadvantage of instability; with ion beam densities of about $1\mu a/cm^2$, about half of the target would disappear in a few minutes of bombardment, but the loss rate thereafter decreased slowly.

Singly-charged He^3 ions were used in the energy region 100 to 400 kev, and doubly-charged ions in the range 320 to 800 kev. In the region of overlap, agreement in yield was observed within an accuracy of 0.5 percent—evidence of lack of He^{3++}-to-He^{3+} conversion through charge neutralization. Likewise there was negligible He^{3+}-to-neutral He^3 conversion, as shown by a test in which the charged beam was deflected from the target by means of an external magnet.

To keep up with the loss of target material, a yield determination was made at 640 kev, followed by one at the energy under investigation, and by a repeated determination at 640 kev. Above 300 kev, the two 640 kev yields agreed within 3 percent, but at 100 kev the target loss between the two 640-kev determinations was as much as 30 percent. The yield at the energy under investigation was always related to the mean of the two 640-kev determinations.

The second part of the experiment consisted of a comparison of the peak yields of the He^3 reaction and the tritium reaction. This was accomplished using the arrangement of Fig. 1, except that only the alpha-particle detector (proportional counter) was used. The target requirement now was one of stability rather than thinness, so a zirconium-deuterium target 40 $\mu g/cm^2$ thick was installed. (At the peaks, the yields vary only slowly with ion energy, so the requirement of thinness could be relaxed.) The target had been made by evaporation of the required thickness of zirconium onto a platinum backing, followed by heating in a deuterium atmosphere. Using doubly-charged ions, the peak in the yield curve was located and carefully examined with the ion beam slightly defocused so as to reduce errors arising from non-uniform coverage of a possibly non-uniform target. The ion source of the Cockcroft-Walton accelerator and its gas supply were then removed and replaced by similar equipment containing tritium. With target and detector unchanged, alpha-particle yields were determined at the peak of the tritium reaction cross section curve. Then the He^3 ion source was replaced, and the peak yield of the He^3 reaction was redetermined; no change in yield was observed, so it was assumed that the target and detector had remained constant, and the result of the intervening tritium bombardment was taken as valid.

The gas used in the tritium bombardment just described contained 25 percent H^1. The possible presence of H_3^+ ions was guarded against by using low pressures in the ion source, and their absence was checked by use of the mass 4 beam (HT^+). The mass 4 beam was found to change with time, possibly because of an admixture of oxygen which cleaned up slowly as the system was

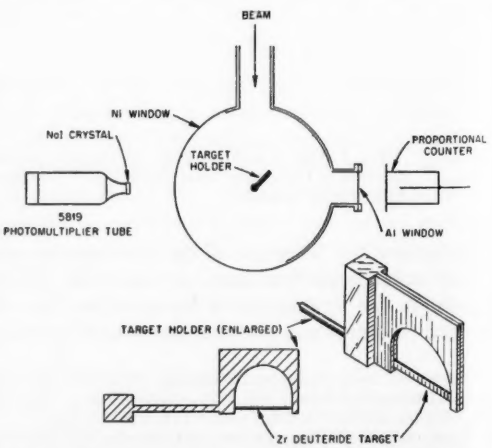


FIG. 2. The apparatus used to determine the angular distribution of the protons from the He^3 reaction. The NaI crystal-photomultiplier combination could be set at various angles to the beam. The fixed proportional counter served as a monitor for the protons produced by the reaction.

[14] 0.5 percent of the total He^3 ion beam.

degassed through operation. (O^+, if accelerated and subsequently stripped to O^{++}, would be magnetically deflected as HT^+.) However, after several hours of operation, the yields from the mass 3 and mass 4 beams agreed within a few tenths of one percent, which implied very little H_3^+ contamination.

A third part of the experiment was required because the He^3-H^3 comparison was carried out at only one angle, *viz.*, 90 degrees relative to the direction of the ion beam. The products of the tritium reaction are known to be isotropic in the center-of-mass system, but similar information was lacking for the He^3 reaction at these energies. Figure 2 shows the target and detector arrangement used to survey the angular distribution. The target chamber was fitted with an extensive nickel window 0.012 cm thick, and a moveable NaI crystal served to count the reaction protons in various directions. The proportional counter at the right counted protons at a fixed angle and served as a monitor. The thick target was a strip 3 mm wide of deuterated zirconium foil, 0.005 cm thick, placed at the axis of rotation of the NaI crystal 20 cm away. In a test of the energy sensitivity of the efficiencies of the two proton counters, the target was rotated 90 degrees without affecting the relative counting rates, whence it was concluded that the target was thin so far as the penetration effects for the protons were concerned.

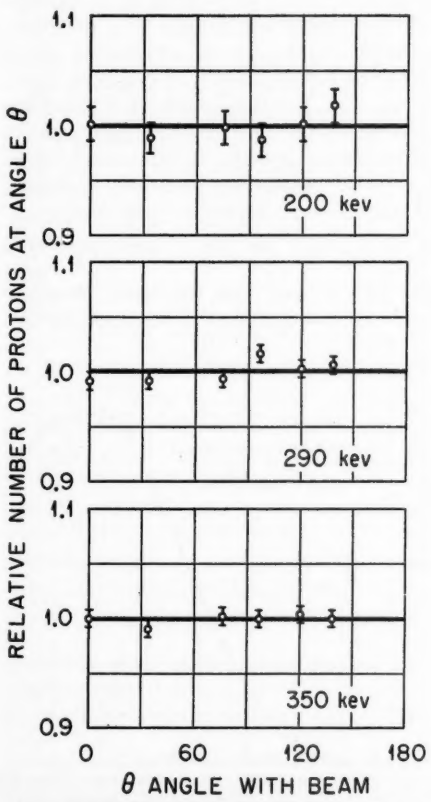


FIG. 3. The center-of-mass angular distribution of the protons from the He^3 reaction. A thick Zr-deuterium target was bombarded at the three energies given. Errors shown are the standard deviations.

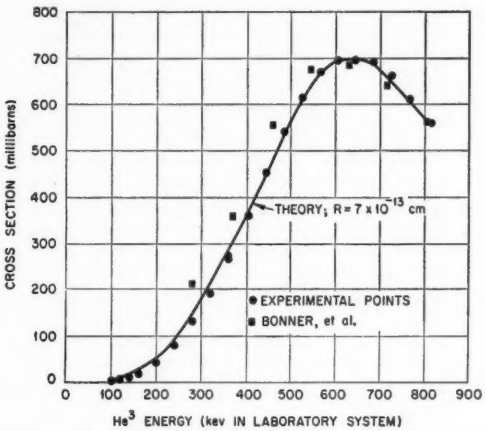


FIG. 4. The cross section of the He^3 reaction as a function of He^3 energy. The solid line is the theoretical cross section obtained by fitting the experimental points by the dispersion formula with the interaction radius, $R = 7 \times 10^{-13}$ cm. The circles are the cross section values determined by this experiment and the squares are the experimental values of Bonner *et al.*

RESULTS

The results of the angular distribution measurements are indicated in Fig. 3. The measurements were taken at 200, 290, and 350 kev and indicate isotropy within limits of about 2 percent; half of the error comes from geometrical uncertainty and half from counting statistics. Earlier results of Bonner *et al.*[10] and Yarnell *et al.*[11] indicate isotropy at somewhat higher energies. It is concluded that at these energies the relative yields of the H^3 and He^3 reactions taken at a single angle can be used directly in deriving the relative reaction cross sections.

Taking the value 5.00 barns (Conner *et al.*[4]) for the cross section at the peak of the tritium reaction, the value 695 millibarns is obtained for the peak cross section of the He^3 reaction. With this normalization, the curves of Figs. 4 and 5 (details of low-energy region) have been derived to show the variation of the cross section of the He^3 reaction with energy. The value 695 millibarns for the cross section at the peak is in agreement with the separate determination by Bonner *et al.*,[10] but it is considerably lower than the 900 millibarns obtained by Yarnell *et al.*[11] and the 940 millibarns obtained by Frier and Holmgren.[13] In Fig. 5, the comparison with the results of Tuck *et al.*[9] indicates lack of agreement amounting to 5 to 10 kev in the energy scale. However, the result of Jarvis and Roaf[12] at 65.6 kev equivalent He^3 energy also indicates a displacement of 6 kev toward higher values than given by Tuck, and is in fact in good agreement with the present results.

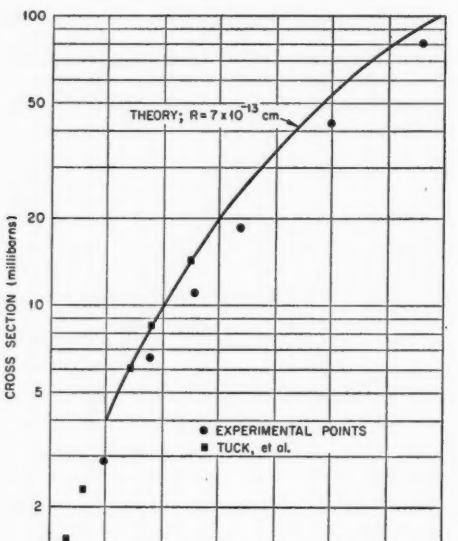


FIG. 5. The expansion of the low energy region of the He^3 reaction cross section. The solid line is the expansion of the theoretical curve of Fig. 4. The circles are the values of the cross section determined by this experiment. The squares are the experimental values of Tuck *et al.*

The experimental error involved in the yield comparison at the peaks of the tritium and He^3 reactions is estimated at less than two percent. Included in this estimate is the spread of about 1 percent in several counting rate comparisons taken with the He^3 and H^3 beams as described above, with slightly varying beam diameters. The current integrator was considered accurate within 0.2 percent, while charge exchange in the H^3 and He^3 beams might amount to a few tenths of one percent as judged from the observations of Kanner[15] and Snitzer.[16] Shifting of the beam on the target might cause geometrical errors which are hard to estimate, but the defocusing of the beam should have reduced this effect far below the maximum value of 6 percent which it would have had in the case of maximum possible shift of a point focus. These considerations taken together led to the cited 2 percent estimated accuracy in the ratios of the peak cross sections.

The cross sections derived for the remainder of the curve have in them the errors involved through the instability of the polyethylene target. As mentioned above, these should be negligible above 400 kev, in view of the manner in which the data were taken. At 100 kev, however, where the target loss was as much as 30 percent in each determination, errors of as much as 5 percent are conceivable, arising from this cause. Changes in the target-detector solid angle caused by beam shift should be negligible because the NaI crystal face was parallel to the target plane. Above 400 kev, where the He^{++} beam was used, defocusing was impossible because the beam was too weak, so about 2 percent error might be involved because of nonuniformity of the thickness of the polyethylene.

In summary, apart from the error in the reference value of 5.00 barns at the peak of the tritium reaction, it is thought that the cross section data of Figs. 4, 5, 6, and 7 are correct in absolute value within ± 10 percent from 100 kev to 400 kev, and ± 5 percent from 400 to 800 kev.

DISCUSSION

The isotropy of the He^3 reaction, herein observed, matches that previously found for the tritium reaction in the same energy region (0 to 200 kev). Because nuclear forces appear to be charge symmetric, the interchange of neutrons and protons in the compound nucleus should affect the Coulomb energy only, and one would expect the nuclear states involved in these two reactions to be identical so far as angular momentum and parity are concerned. States giving isotropy are those with either $J=\frac{1}{2}$ or $l=0$. In both reactions the spin s of the bombarding particle is $\frac{1}{2}$, and the spin i of the target nucleus is one. Therefore the initial states must be either doublets or quartets: ^2S, ^4S, ^2P, ^4P, ^2D, ^4D. . . . The final states of the systems are characteristic of an alpha particle (spin zero) and a neutron or proton (spin $\frac{1}{2}$), so they will be doublets: ^2S, ^2P, ^2D. . . . Assuming spin-orbit coupling, the allowed transitions giving isotropy are then

$$^2S_{\frac{1}{2}} \rightarrow {}^2S_{\frac{1}{2}}, \quad {}^4S_{\frac{1}{2}} \rightarrow {}^2D_{\frac{1}{2}}, \quad {}^2P_{\frac{1}{2}} \rightarrow {}^2P_{\frac{1}{2}}, \quad {}^4P_{\frac{1}{2}} \rightarrow {}^2P_{\frac{1}{2}}, \quad {}^4D_{\frac{1}{2}} \rightarrow {}^2S_{\frac{1}{2}}.$$

We shall now assume that the same transitions are involved in the two reactions. If we assume further that

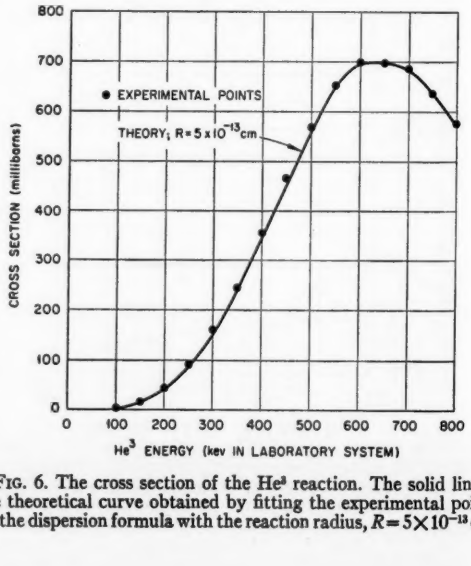


FIG. 6. The cross section of the He^3 reaction. The solid line is the theoretical curve obtained by fitting the experimental points by the dispersion formula with the reaction radius, $R=5\times10^{-13}$ cm.

[15] H. Kanner, Phys. Rev. **84**, 1211 (1951).
[16] E. Snitzer, Phys. Rev. **89**, 1237 (1953).

only one J value is involved, we can deduce whether it is $\frac{1}{2}$ or $\frac{3}{2}$ from considerations of the value of the maximum cross section.[17] The maximum reaction cross section σ_r is given by the expression

$$\sigma_r=\frac{2J+1}{(2s+1)(2i+1)}\pi\lambda^2,$$

so if $J=\frac{1}{2}$, the maximum value that the cross section can have is $\pi\lambda^2/3$. Actually the experimental value for the cross section of the tritium reaction at 200 kev is twice this, so the value $J=\frac{3}{2}$ is required, and the transition must be $^4S_{\frac{1}{2}}\rightarrow{}^2D_{\frac{3}{2}}$. Thus if the same *single* compound state is involved in the two reactions, it must have $J=\frac{3}{2}$ and be of even parity, and we must have $l=0$ for the incident wave. These J and l values will be adopted in the following discussion.

Argo *et al.*[6] were able to fit the experimental shape of the tritium reaction resonance with the level shift form of the one-level dispersion formula. In the notation of Blatt and Weisskopf,[18] the reaction cross section σ_r is given by

$$\sigma_r=\frac{2J+1}{(2s+1)(2i+1)}\pi\lambda^2\frac{\Gamma_\alpha\Gamma_r}{[\epsilon-\epsilon_s+\Delta_l(\epsilon)\gamma_\alpha]^2+\frac{1}{4}(\Gamma_\alpha+\Gamma_r)^2},$$

where $2\pi\lambda$ is the de Broglie wavelength associated with the reduced mass of the particles in their relative motion, ϵ is the center-of-mass energy of the incident particle, ϵ_s is that value of ϵ for which the logarithmic derivative of the wave function is zero and is the "formal" resonance energy, Γ_α and Γ_r are the scattering and reaction widths respectively, γ_α and γ_r are the reduced scattering and reaction widths, and $\Delta_l(\epsilon)$ is given by

$$\Delta_l(\epsilon)=\frac{R}{F_l^2+G_l^2}\left[F_l\frac{dF_l}{dr}-G_l\frac{dG_l}{dr}\right]_{r=R},$$

where in turn R is the radius of interaction and F_l and G_l are respectively the regular and irregular solutions of the radial wave equation as tabulated by Bloch *et al.*[19] The reduced scattering and reaction widths, γ_α and γ_r, are designed to be almost independent of Coulomb effects, and are determined through the expressions

$$\gamma_\alpha=\Gamma_\alpha(F_l^2+G_l^2)/2kR,$$
$$\gamma_r=\Gamma_r/2k_rRP_r,$$

where k is the wave number of the incident particle, k_r is that of the emitted particle, and P_r is the penetrability factor associated with the emitted particle. Using $R=7\times10^{-13}$ cm, Argo *et al.*[6] found two sets of level parameters yielding an energy variation of the cross section which agreed with their experiments. The two sets of parameters were as follows:

ϵ_s(kev)	γ_{H^3}(kev)	γ_n(kev)
−67	471	15.7
−126	715	17.2

Following the same line of reasoning and with $R=7\times10^{-13}$ cm we have analyzed[20] our results for the He^3 reaction using an iterative least squares method on IBM machines,[21] and we likewise find two sets of fitting parameters:

ϵ_s(kev)	γ_{He^3}(kev)	γ_p(kev)
235	62	23
129	780	12

Figures 4 and 5 show the comparison of the theoretical curve, thus calculated, with the experimental points. If we now select the second set of parameters in each case, we find quite reasonable agreement:

$$\gamma_{He^3}/\gamma_{H^3}=1.1, \quad \text{and} \quad \gamma_p/\gamma_n=0.7.$$

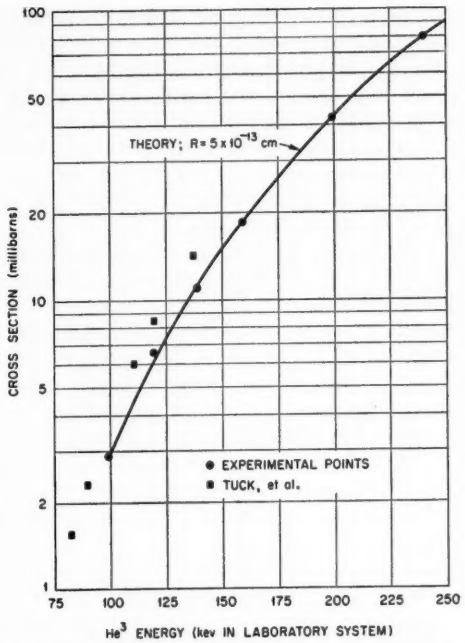


FIG. 7. The expansion of the low energy region of the He^3 reaction cross section. The solid line is the expansion of the theoretical curve of Fig. 6. The circles are the values of the cross section determined by this experiment. The squares are the experimental values of Tuck *et al.*

[17] A. Simon and T. A. Welton, "A Note on the Maximum Cross Section for Resonance Reactions," Oak Ridge National Laboratory Memorandum, (unpublished).

[18] J. M. Blatt and V. E. Weiskopf, *Theoretical Nuclear Physics* (John Wiley and Sons, Inc., New York, 1952), Chap. VIII.

[19] Bloch, Hull, Broyles, Bouricius, Freeman, and Breit, Revs. Modern Phys. 23, 147 (1951).

[20] Actually the fit for $R=7\times10^{-13}$ cm was to the nonlevel-shift form of the dispersion formula and was converted to the above form by the linear approximation of $\Delta_l(\epsilon)$ as outlined by R. G. Thomas, Phys. Rev. 81, 148 (1951).

[21] We are indebted to the computer staff of the K-25 laboratory for this assistance.

The errors in these ratios are difficult to estimate because systematic errors in the cross section measurement in the low-energy region would introduce rather large errors in the reduced widths. It is felt that the reduced widths of the He^3 reaction are within ± 25 percent of the values given. If the errors in the tritium reaction measurements are comparable, then the reduced widths associated with the two reactions would agree within experimental error.

The positions of the resonance levels in He^5 and Li^5 are given by $Q+\epsilon_s-$ (the breakup energy into He^4 plus a neutron or a proton, respectively). For He^5, the Q value is 17.577 Mev, and the breakup energy is 1.0 ± 0.1 Mev so that the level is $17.577-0.126-1.0\pm 0.1 = 16.5\pm 0.1$ Mev above the ground state. Similarly the position of the level in Li^5 is $18.341+0.129-1.8\pm 0.3 = 16.7\pm 0.3$ Mev above the ground state. Thus the level positions, as well as the reduced widths, agree within experimental error. The agreement indicates that the difference between the two reactions consists mainly of Coulomb effects.

Let us now examine the effect of varying the parameter R. Following Conner *et al.*[4] and Bonner *et al.*,[10] we choose the value $R=5\times 10^{-13}$ cm. In this case, those investigators found violent disagreement between the resonance parameters of the two reactions, in that the value of the reduced scattering width of the tritium reaction turned out to be fourteen times larger than that of the He^3 reaction. We have taken our new results for the He^3 reaction and have found, as was the case for $R=7\times 10^{-13}$ cm, *two sets* of fitting parameters:

γ_p(kev)	γ_{He^3}(kev)	ϵ_s(kev)
36.3	116	205
41.9	2930	−391

Figures 6 and 7 illustrate the agreement between the experimental points of the theoretical cross section curve computed by the use of either of these sets; the fit is very satisfactory.[22] Now if we choose the second set for comparison with those of Conner *et al.*[4] for the tritium reaction, we obtain

$$\gamma_{He^3}/\gamma_{H^3}=2930/2000=1.5,$$
$$\gamma_p/\gamma_n=41.9/56=0.8.$$

Allowing as before about ± 25 percent experimental uncertainty, it is apparent that on this basis the reduced widths of the resonances in the two reactions are substantially equal. The apparent discrepancy noted in the Rice Institute work resulted from their use of only one set of fitting parameters for the He^3 reaction. The position of the level in He^5 is found to be 16.1 ± 0.1 Mev above the ground level, and in Li^5 the level is at 16.2 ± 0.3 Mev above ground.

Using $R=5\times 10^{-13}$ cm, two points for the He^3 reaction cross section were calculated for higher energies, one at 1190-kev and the other at 1890-kev He^3 energy. The calculated value was 85 percent of the experimental value at 1190 kev and 50 percent of the experimental value at 1890 kev. The experimental values of Bonner *et al.*[10] were used. This comparison, along with the fact that the angular distribution becomes anisotropic at energies above the maximum of the cross section curve, shows that the one-level dispersion formula loses validity rather quickly beyond the maximum of the resonance.

We conclude that both the tritium and the He^3 reactions can be fitted over a large part of the resonance region by the one-level dispersion formula and that the corresponding reduced widths so obtained are equal within the experimental uncertainties. The fact of equality between the reaction widths is rather insensitive to the value chosen for the reaction radius. If ϵ_s is used to obtain the level position in the compound nuclei, He^5 and Li^5, then these levels correspond to the same excitation in the compound nuclei within the rather large experimental uncertainty.

Since nuclear forces are known to be charge-symmetric, the agreement between the two sets of reaction parameters was to be expected if the dispersion formula treatment is valid and if Coulomb effects can be separated from the effects arising from nuclear forces. This indeed is the case in the region near the top of the resonance and at lower energies, but at higher energies the one-level dispersion formula treatment does not agree with experiment.

ACKNOWLEDGMENTS

The author wishes to thank Dr. W. M. Good for suggesting the problem and for a very large amount of advice and assistance with all phases of the work. The author also wishes to thank Dr. R. D. Present and Dr. A. Simon for their assistance with the theory, and Dr. H. E. Banta, Dr. R. L. Macklin, and Mr. J. B. Dial for their help with the preparation of targets. He is also indebted to Dr. J. Moshman who did the statistical analysis with the aid of Mrs. Joan Vander Sluis, and to Mr. V. Anderson, who performed the IBM calculations. To Dr. A. H. Snell is extended grateful acknowledgment for advice and criticism in the preparation of this manuscript.

[22] In this case, the fit was to the level-shift form of the dispersion formula. Thus errors introduced by the linear approximation method used for $R=7\times 10^{-13}$ cm were eliminated. This accounts for the better fit to the experimental points.

β Spectrum of Bi^{210}(RaE) and the Coupling Constants of Scalar and Tensor Interactions in β Decay

G. E. LEE-WHITING
Chalk River Laboratory, Atomic Energy of Canada Limited, Chalk River, Ontario, Canada
(Received August 25, 1954)

It is shown that the only pure states of spin 1 having magnetic moments in agreement with the measured upper limit for Bi^{210} are $(h_{9/2}; g_{9/2})_{1-}$ and $(h_{9/2}; i_{11/2})_{1-}$. The former is inconsistent with the results of an analysis of the β spectrum of Bi^{210} with a linear combination of scalar and tensor interactions, while the latter is consistent. It follows from the acceptance of the correctness of the latter that the relative sign of the scalar and tensor coupling constants is positive.

1. INTRODUCTION

VARIOUS authors[1-3] have now shown that it is extremely difficult to explain the non-allowed shape of the β spectrum of Bi^{210} with a mixture of tensor and pseudoscalar interactions and with the assumption that the transition is 0− to 0+. On the other hand Yamada[4] has demonstrated that it is possible to obtain this shape, with a mixture of scalar and tensor interactions, for the transition 1− to 0+. It thus appears that the decay of Bi^{210} is a case of a first forbidden transition with non-allowed shape of the type 1− to 0+. From a fitting of the theoretical to the observed correction factor it is possible to obtain relations involving the coupling constants of the interactions, G_S and G_T, and the various nuclear matrix elements concerned. For most nuclei it is extremely difficult to learn much about G_S and G_T from these relations because little is known about the relative magnitudes of the nuclear matrix elements. However, the ratio of $\int\beta\sigma\times\mathbf{r}$ to $\int\beta\mathbf{r}$, one of the quantities needed for first forbidden transitions, can be calculated very easily if the initial and final states are spectroscopically pure. The nucleus Bi^{210}, having but one nucleon of each kind apart from closed shells, is one of the nuclei most likely to satisfy this condition. In addition, it is believed that for values of atomic number as high as that of bismuth jj coupling is a good approximation. We shall proceed, then, to find out what can be learned about G_S and G_T when it is assumed that the ground states of Bi^{210} and Po^{210} are spectroscopically pure.

Before one can calculate ratios of nuclear matrix elements it is necessary to identify the configurations concerned. The most useful piece of experimental information for this purpose is the demonstration, by measurement of hyperfine structure, that the magnetic moment of Bi^{210} is exceedingly small.[5] King and Peaslee[6] have shown that the state $(h_{9/2}; g_{9/2})_{1-}$ satisfies this condition. A great objection to their choice, however, is that it is expected to lie above the state of spin 0 of the same configuration. Pryce[7] has calculated the interaction energy, as a function of resultant angular momentum, for jj-coupled neutron and proton, with the assumption of forces of infinitesimal range. Such calculations are very conveniently displayed on a diagram of the sort used by de-Shalit.[8] In this diagram the ordinate is energy and the abscissa, α, is a measure of the strength of the spin-spin forces in the interaction, attractive or repulsive as η is $+1$ or -1; reasonable values of α are believed to lie in the interval 0 to 0.25. The de-Shalit diagram for the configuration $(h_{9/2}; g_{9/2})$ is shown in Fig. 1(a). For all plausible values of α, the state of spin 0 is far below that of spin 1.

In Sec. 2 the magnetic moments of all reasonable ground states for Bi^{210} are calculated. Apart from $(h_{9/2}; g_{9/2})_{1-}$, the only other acceptable choice is $(h_{9/2}; i_{11/2})_{1-}$. The de-Shalit diagram for this configuration is shown in Fig. 1(b). For most reasonable values of α the state of spin 1 is lowest. The low-lying state of spin 10 may well be the long-lived α-emitting isomer of Bi^{210} found by Neumann *et al.*[9]

In Sec. 3 the shape of the β spectrum of Bi^{210} is considered. This shape is explicable with the ratios of matrix elements calculated for the assumption that $(h_{9/2}; i_{11/2})_{1-}$ is the ground state of Bi^{210}, but not for those ratios corresponding to the assumption of $(h_{9/2}; g_{9/2})_{1-}$ for this ground state. Acceptance of the first choice requires that G_S/G_T be positive.

2. MAGNETIC MOMENTS

The magnetic moment of a system consisting of a proton and a neutron whose angular momenta, j_1 and j_2, are coupled to a resultant J is desired. The orbital angular momenta of the proton and neutron will be represented by l_1 and l_2 respectively. By the methods of Racah[10] it is possible to show that the required

[1] H. Takebe, Progr. Theoret. Phys. (Japan) **10**, 673 (1953).
[2] M. E. Rose and R. K. Osborn, Phys. Rev. **93**, 1315 (1954).
[3] Alaga, Kofoed-Hansen, and Winther, Kgl. Danske Videnskab. Selskab, Mat.-fys. Medd, **28**, No. 3 (1953).
[4] M. Yamada, Progr. Theoret. Phys. (Japan) **10**, 252 (1953).
[5] Fred, Tomkins, and Barnes, Phys. Rev. **92**, 1324 (1953).
[6] R. W. King and D. C. Peaslee, Phys. Rev. **94**, 795 (1954).
[7] M. H. L. Pryce, Proc. Phys. Soc. (London) **A65**, 773 (1952).
[8] A. de-Shalit, Phys. Rev. **91**, 1479 (1953).
[9] Neumann, Howland, and Perlman, Phys. Rev. **77**, 720 (1950).
[10] G. Racah, Phys. Rev. **62**, 438 (1942).

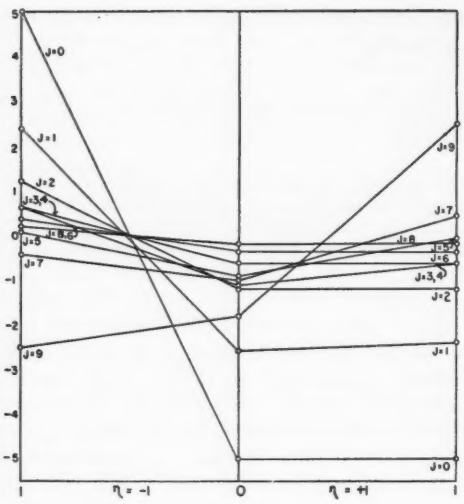


FIG. 1. De-Shalit diagrams (a) for the configuration $(h_{9/2}; g_{9/2})$; (b) for the configuration $(h_{9/2}; i_{11/2})$.

magnetic moment is

$$\mu=\frac{1}{J+1}\left[\frac{J(J+1)+j_1(j_1+1)-j_2(j_2+1)}{2j_1}\mu_p(l_1j_1)+\frac{J(J+1)+j_2(j_2+1)-j_1(j_1+1)}{2j_2}\mu_n(l_2j_2)\right],$$

where $\mu_p(l_1j_1)$ and $\mu_n(l_2j_2)$ are the magnetic moments, in the single-particle states indicated, of a proton and of a neutron respectively. The "Schmidt model" values of μ_p and μ_n will be used.[11]

The question of the choice of single-particle states has been discussed by Pryce;[7] several choices are almost equally probable. For the proton the shell-model suggests $h_{9/2}$, $f_{7/2}$, or possibly $i_{13/2}$; support for $h_{9/2}$ comes from the measured spin and magnetic moment of Bi^{209}. For the neutron we have $g_{9/2}$, $i_{11/2}$, or perhaps $k_{15/2}$; no spins of even-odd nuclei in this region are known. In Table I the magnetic moments of all states of spin 1 or 2 and of either parity which may be formed from the suggested single-particle states are listed.

Because the magnetic moments of all states of even parity are large, the ground state of Bi^{210} cannot have even parity. The assignment 1+ would require a β spectrum of allowed shape, contrary to observation, but the assignment 2+, with a second forbidden β-transition, was at one time thought to explain the spectrum shape.[12] The only states with small magnetic moments are those with an $h_{9/2}$ proton; all other states of odd parity have magnetic moments too large to be considered. Because of the well-known discrepancy between calculated and observed nuclear magnetic moments, none of the cases in the first column can be excluded, even if the magnetic moment is a little larger than the observed upper limit. Fortunately the states of spin 2− may be rejected because their β spectra, which have unique shape, are not in agreement with experiment. Only the two states of spin 1 in the first column need be given further consideration.

3. β SPECTRUM OF Bi^{210}

The experimental shape of the spectrum which we shall use is that given by Langer and Price.[13] From their curve the experimental correction factor was obtained; the maximum electron energy used was 3.29 relativistic units. Expressions for the direct and cross terms of the theoretical correction factor have been given by Greuling[14] and Pursey[15] respectively. If it is assumed that the interaction is of the form $G_S S+G_T T$, then in the correction factor there appear three different matrix elements, among which we write the following relations:

$$\int \beta\alpha = \Lambda(\alpha Z/2\rho)\int \beta\sigma\times\mathbf{r} \quad \text{and} \quad \int \beta\sigma\times\mathbf{r} = i\epsilon\int \beta\mathbf{r}.$$

It is convenient to put $x=G_S/(G_T\epsilon)$. We shall assume that Λ and x are real; justification of this assumption

[11] J. M. Blatt and V. F. Weisskopf, *Theoretical Nuclear Physics* (John Wiley and Sons, Inc., New York, 1952).
[12] E. J. Konopinski and G. E. Uhlenbeck, Phys. Rev. **60**, 308 (1941).
[13] L. M. Langer and H. C. Price, Phys. Rev. **76**, 641 (1949).
[14] E. Greuling, Phys. Rev. **61**, 568 (1942).
[15] D. L. Pursey, Phil. Mag. **42**, 1193 (1951).

has been given by Longmire and Messiah.[16] For ease in handling the cancellation of the large terms in the correction factor, it is convenient to examine its dependence on the two variables x and y, y being defined by the relations $y=\frac{1}{2}(1+S)\Lambda-1-x$ and $S=[1-(\alpha Z)^2]^{\frac{1}{2}}$. The correction factor C for the transition 1− to 0+ is given by expression (1).

$$C=2G_T^2(1+S)^{-1}(\alpha Z/2\rho)^2 D\epsilon^2\left|\int \mathbf{r}\right|^2, \qquad (1)$$

where

$$D=y^2+\zeta\{-2y[(2S+1)^{-1}(1+x)(W-S^2W^{-1})+(1+S)(1-x)K/6]-(1-S)(2S+1)^{-1}y^2\times[(2S+3)W+SW^{-1}]\}+\tfrac{1}{2}(1+S)\zeta^2\times[(F_1-F_3+F_5)x^2+(2F_5-F_3)xy+F_5y^2+(2m_0-F_3+F_4+2F_5)x+(F_4+2F_5)y+(F_2+F_4+F_5)+2(x-\tfrac{1}{2})^2L_1],$$

$$F_1=(1+S)K^2/6-\tfrac{2}{3}(2S+1)^{-1}\times K(Sp^2W^{-1}-2\alpha^2Z^2W)+m_0,$$

$$F_2=(1+S)K^2/12+\tfrac{2}{3}(2S+1)^{-1}\times K(Sp^2W^{-1}-2\alpha^2Z^2W)+m_0,$$

$$F_{3,4}=2\alpha Z(1+S)^{-1}\{\tfrac{1}{3}K\alpha Z(2S+1)^{-1}\times[(2S+3)W+SW^{-1}]\mp n_0,$$

$$F_5=(1-S)(1+S)^{-1}l_0, \quad \zeta=2\rho/\alpha Z.$$

For the definitions of L_0, M_0, N_0, L_1, and K see reference 12; m_0, n_0 and l_0 are the coefficients of the zeroth, first and second powers of ρ in the expansions of M_0, N_0, and L_0 (see reference 4). Values of L_1 were taken from the tables of Rose *et al.*[17] The expression for D is correct as far as the term in ρ^2. For $x=\frac{1}{2}$ expression (1) reduces to Yamada's Eq. (13), except for some small terms in ρ^2. In discussing the shape of the spectrum one need only consider D, the other factors in C being independent of energy.

We next ask what values of x and y correspond to correction factors of the observed form. To answer this question we equate the theoretical and experimental ratios of the correction factors at $W=1.2$ and $W=2$. Since D is a quadratic function of x and y, any such relation between the values of the correction factor at two different energies must give a conic section in the x-y plane. The points in the x-y plane satisfying this particular equation lie on an hyperbola. Because there is some uncertainty in the measured ratio, we have used two values, 1.57 and 1.61. The curves labelled 1.57 and 1.61 in Fig. 2 are the corresponding hyperbolas. If a point (x,y) is to correspond to a correction factor whose value at $W=1.2$ is in correct ratio to its value at $W=2$, it must lie in the region bounded by the two hyperbolas. A second condition on x and y may be obtained by equating theoretical and observed values of the ratio of the correction factor at $W=3$ and $W=2$. If the point (x,y) is to correspond to a correction factor whose value at $W=3$ lies between 0.49 and 0.55 times the value at $W=2$, then this point must lie between the two ellipses labelled 0.49 and 0.55. To satisfy both conditions a point must lie in the hatched region of Fig. 2. Not all points in this region need, however, correspond to good correction factors; the correction factor must be tested over the whole range of W.

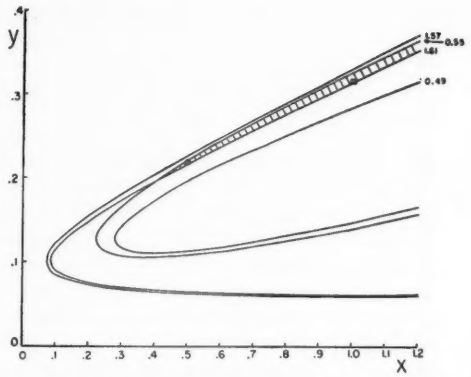


FIG. 2. The hatched area is the region of the x-y plane in which the theoretical and observed correction factors are in agreement. At any point between the hyperbolas labelled 1.57 and 1.61 the ratio of the values of the correction factor at $W=1.2mc^2$ and $2mc^2$ is in agreement with experiment; between the ellipses marked 0.49 and 0.55 the ratio of the values of the correction factor at $3mc^2$ and $2mc^2$ is in agreement with experiment. The point marked ○ corresponds to the correction factor found by Yamada, (see reference 4) that marked □ to the one found by the author.

The fit found by Yamada[4] corresponds to the point $(x=\frac{1}{2}, y=0.219)$; the corresponding correction factor is shown in Fig. 3. We have obtained an equally good correction factor for the point (1,0.316), also shown in Fig. 3. A reasonable fit is probably possible anywhere in the hatched area.

In calculating correction factors the nuclear radius was assumed to be 8.3×10^{-13} cm. A slightly different choice for this radius would not affect the positions of the hyperbolae and ellipses of Fig. 2 very much; in particular, we have shown that when ρ lies between 4.7×10^{-13} cm and 10.7×10^{-13} cm the minimum value

TABLE I. Magnetic moments of possible *jj*-coupled ground states of Bi^{210}. In each compartment of the table are inserted the spins, parities and calculated magnetic moments of certain states formable from the single-particle states indicated.

Neutrons \ Protons	$h_{9/2}$	$f_{7/2}$	$i_{13/2}$
$g_{9/2}$	1−, 0.080 2−, 0.16	1−, −4.07 2−, −1.89	2+, 8.04
$i_{11/2}$	1−, −0.36 2−, +0.34	2−, −2.59	1+, 4.26 2+, 3.94
$k_{15/2}$			1−, −5.48 2−, −2.92

[16] C. L. Longmire and A. M. L. Messiah, Phys. Rev. **83**, 464 (1951).

[17] Rose, Perry, and Dismuke, Oak Ridge Laboratory Report ORNL-1459, 1953 (unpublished).

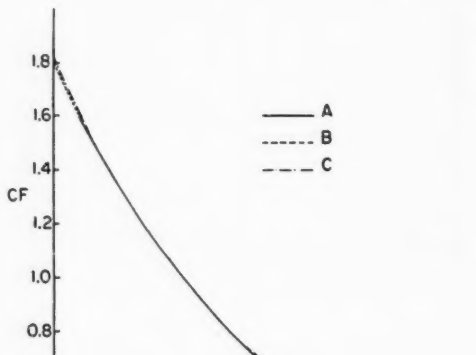


FIG. 3. Correction factor for the spectrum of Bi^{210}. The experimental curve is referred to as *A*, the fit found by Yamada as *B*, that by the author as *C*.

of x attained by the ellipse labelled 0.55 is greater than 0.19.

4. DISCUSSION

Examination of Fig. 2 leads to the conclusion that the theoretical correction factor will not explain the observed spectrum unless x is greater than 0.35. This limit will now be used to show that the assignment of $(h_{9/2}; g_{9/2})_{1-}$ to the ground state of Bi^{210} is inconsistent with its spectrum. Recall that $x=G_S/(G_T\epsilon)$. We are assured by the work of Kofoed-Hansen and Winther[18] and by that of Blatt[19] that $|G_S/G_T|$ is not appreciably greater than unity. It has been proved by Pursey[15] and by Ahrens and Feenberg[20] that

$$\epsilon=j_p(j_p+1)-j_n(j_n+1)-l_p(l_p+1)+l_n(l_n+1).$$

In this formula l and j are the orbital and total angular momentum quantum numbers of a single-particle state; the subscripts p and n refer to proton and neutron respectively. Since for the configuration $(h_{9/2}; g_{9/2})$ the value of ϵ is -10, the corresponding value of x cannot possibly be as large as 0.35, and therefore this choice of configuration cannot explain the β spectrum. For the assumption $(h_{9/2}; i_{11/2})$, ϵ is $+1$, and the observed spectrum is obtainable for a wide range of values of G_S/G_T above 0.35. Thus, if the ground state of Bi^{210} is a pure state, then that pure state must be $(h_{9/2}; i_{11/2})_{1-}$; the state required by the spectrum analysis is the same as that preferred by the energy-level calculation.

A comparison of values of Λ predicted with those found by spectrum-fitting is useful. Rose and Osborn[21] expect Λ to lie in the range 1 to 3, with upper value considered somewhat more plausible. Pursey[15] and Ahrens and Feenberg[20] came to a similar conclusion. The value of Λ obtained from fitting the spectrum of Bi^{210} depends on the value of x assumed; for $x=\frac{1}{2}$ we find $\Lambda=1.92$ and for $x=1$ we find $\Lambda=2.59$. Since the two values of Λ are in equally good agreement with prediction, it is not possible to use them to determine G_S/G_T. On the other hand, the agreement suggests that the fit found for the spectrum is the correct one.

The positiveness of the ratio G_S/G_T, which follows from the positiveness of x and ϵ, is not in agreement with earlier determinations.[22,23] Insofar as the work of Peaslee is based on the spectrum of Bi^{210} it is invalidated by the work of Rose and Osborn.[2] We find too that the first forbidden correction factor for $G_S S+G_T T$ used by this author does not agree with expression (1). The Japanese authors reach the same conclusion as Peaslee from a study of almost the same decaying nuclei, but without finding it necessary to correct the nuclear matrix elements for the presence of pseudoscalar forces. Of the four nuclei which Morita *et al.* consider, only Cs^{137} definitely requires a negative value of G_S/G_T. We believe that ratios of nuclear matrix elements calculated with pure jj-coupled wave functions are more likely to be correct for Bi^{210} than for Cs^{137}. If it is accepted as certain that Λ must be positive, then one of these four spectra, that of Fe^{59}, actually requires a positive value for the ratio G_S/G_T.

The identification of the configuration of the ground state of Bi^{210} enables one to make other shell-model assignments. The transition Pb^{209} to Bi^{209} has logft equal to 5.51, calculated for a maximum electron energy of 0.620 Mev[24] with the aid of tables of the Fermi function.[25] A first forbidden transition with such a small ft-value cannot correspond to cancellation of terms in the correction factor as in Bi^{210}. Such cancellation, if we accept the assignment $(h_{9/2}; i_{11/2})_{1-}$ to Bi^{210}, occurs for $\epsilon=+1$. The ground state of Bi^{209} is known to be $h_{9/2}$. The values of ϵ for the two possible assignments to Pb^{209}, $g_{9/2}$ and $i_{11/2}$, are -10 and $+1$ respectively. The latter must be rejected. This means that the 127th is in different states in Pb^{209} and in Bi^{210}. We expect, therefore, that the $i_{11/2}$ level in Pb^{209} should have a very small excitation energy, and that the radial integral fixing the magnitude of the neutron-proton interaction energy[7] should be larger for the pair $(h_{9/2}; i_{11/2})$ than for the pair $(h_{9/2}; g_{9/2})$. Since the h and i radial wave functions are those which have no nodes and since the

[18] O. Kofoed-Hansen and A. Winther, Phys. Rev. **86**, 428 (1952).
[19] J. M. Blatt, Phys. Rev. **89**, 83 (1953).
[20] T. Ahrens and E. Feenberg, Phys. Rev. **86**, 64 (1952).
[21] M. E. Rose and R. K. Osborn, Phys. Rev. **93**, 1326 (1954).
[22] D. C. Peaslee, Phys. Rev. **91**, 1447 (1953).
[23] Morita, Fujita, and Yamada, Progr. Theoret. Phys. (Japan) **10**, 630 (1953).
[24] Hollander, Perlman, and Seaborg, Revs. Modern Phys. **25**, 469 (1953).
[25] *Tables for the Analysis of Beta Spectra*, Applied Mathematics Series No. 13 (National Bureau of Standards, Washington, D. C., 1952).

g has a node, the required behavior of the radial integral is not unreasonable.

Another β transition of interest is that of Pb^{210}(RaD) into Bi^{210}. This decay goes, as far as is known, entirely to an excited state of Bi^{210} at 47 kev. It is consistent with the conclusions of the preceding paragraph to make the assignment $(h_{9/2}; g_{9/2})_{0-}$ to this excited state. The ft-value for this transition should, then, be very near that for the decay of Pb^{209}. Using a maximum electron energy of 18 kev,[26] we find $\log ft = 5.49$. This same assignment will also explain the absence of a transition to the ground state of Bi^{210}, for it requires that the ground state of Pb^{210} be $(\ ; g_{9/2}^2)_0$. A transition from this initial state to the ground state of Bi^{210} would require a two-particle jump.

From a more detailed examination of the ft-values of transitions between nuclei in the neighborhood of Pb^{208} an estimate of the values of Λ and of G_S/G_T can be obtained. Because the ft-values for the transitions $Pb^{209} \rightarrow Bi^{209}$ and $Pb^{210} \rightarrow Bi^{210*}$ are almost equal, we shall equate the expressions for the correction factors corresponding to the two transitions. For the former the correction factor is given by (1) with the addition of the term corresponding to the matrix element $\int \boldsymbol{\sigma}\cdot\mathbf{r}$ occurring in the tensor interaction. The correction factor for this matrix element was given by Konopinski and Uhlenbeck;[12] we shall use only the term of order ρ^{-2}, *viz.*,

$$\frac{2}{1+S}(\alpha Z/2\rho)^2 G_T^2 \left| \int \boldsymbol{\sigma}\cdot\mathbf{r} \right|^2 .$$

It is well known that only the matrix element $\int \boldsymbol{\sigma}\cdot\mathbf{r}$ contributes to transitions of the type $0+ \rightarrow 0-$. As is shown in the appendix, the correction factor corresponding to this matrix element for the two-particle system is exactly twice its value for the same particle jump in the one-particle system. Methods for the calculation of $\int \boldsymbol{\sigma}\cdot\mathbf{r}$ and $\int \mathbf{r}$ in the single-particle system are well known; convenient formulas are given by King and Peaslee.[27] After cancellation of common factors including the radial integrals, the equation obtained by equating the correction factors for the two transitions becomes

$$1+\frac{100}{99}\left(\frac{1+S}{2}\Lambda - 1 + \frac{1}{10}\frac{G_S}{G_T}\right)^2 \simeq 2. \qquad (2)$$

A similar pair of decaying nuclei are Tl^{207} and Tl^{206}, for which the calculated $\log ft$-values are 5.16 and 5.18 respectively. In each case a $p_{\frac{1}{2}}$ neutron changes into an $s_{\frac{1}{2}}$ proton; calculations of the Pryce type indicate that the spin of the ground state of Tl^{206} is $0-$. Equation (3) for the two Tl isotopes corresponds to Eq. (2) for the two Pb isotopes.

$$1+\frac{4}{3}\left(\frac{1+S}{2}\Lambda - 1 - \frac{1}{2}\frac{G_S}{G_T}\right)^2 \simeq 2. \qquad (3)$$

If it may be assumed that Λ has roughly the same value in the two cases, then (2) and (3) may be solved simultaneously. There are four solutions. Two of these require $|G_S/G_T|$ to be greater than 3, contrary to the evidence from allowed transitions. A third may be excluded because of the smallness of the value of Λ, *viz.*, 0.025. For the remaining solution we have $\Lambda = 2.2$ and $G_S/G_T = 0.23$. This value of Λ, which is not very sensitive to the approximations involved, is in good agreement with the values obtained from fitting the spectrum of Bi^{210}. The smallness of the value of G_S/G_T found by this method is of little significance, because the approximations introduce a very large uncertainty in this value; even the sign of the ratio is not really well determined.

5. POSSIBILITY OF CONFIGURATION MIXING

All the arguments and discussion in the preceding parts of this paper are founded on the assumption that the ground state of Bi^{210} is spectroscopically pure. Though there is, we believe, evidence for the assumption, it must be admitted that the possibility that this ground state is a mixture of $(h_{9/2}; i_{11/2})_{1-}$ and $(h_{9/2}; g_{9/2})_{1-}$ cannot be excluded. Such a mixture is not inconsistent with the spectrum shape. Let a and b refer to $(h_{9/2}; i_{11/2})_{1-}$ and $(h_{9/2}; g_{9/2})_{1-}$ respectively; we write $a\psi_a + b\psi_b$ for the wave function of the ground state of Bi^{210}. Instead of having $x = G_S/G_T$ as we did for the pure state ψ_a, we now have $x = (G_S/G_T)F$ for the mixed state. The function F is $(1+t)(1-10t)^{-1}$, where t is the ratio of $b(f|r|b)$ to $a(f|r|a)$; the letter f represents the final state $(h_{9/2}^2; \ )_0$. If we are allowed complate freedom in the choice of t, then F may take any value, and nothing can be said about the ratio G_S/G_T. If, on the other hand, it can be shown that t is negative, then the only negative values for F lie in the range $-1/10$ to 0; since these values are inconsistent with what is known of the magnitudes of x and G_S/G_T, it can then be said that G_S/G_T must be positive. By well known methods the expression for t can be reduced to $(3\sqrt{6})^{-1} bI_1/aI_2$, where

$$I_1 = \int_0^\infty R(1h)R(2g)r^3 dr, \quad I_2 = \int_0^\infty R(1h)R(1i)r^3 dr;$$

the R's stand for normalized radial wave functions. It may also be shown, by consideration of configuration interaction in a system consisting of a jj-coupled proton and neutron, that b/a has the same sign as $-(a|V|b)$, where V is the interaction potential between the proton and the neutron. Using the theory of Pryce,[7] one may

[26] Insch, Balfour, and Curran, Phys. Rev. 85, 805 (1952).
[27] R. W. King and D. C. Peaslee, Phys. Rev. 94, 1284 (1954).

show that $(a|V|b)$ has the same sign as

$$I_3=\int_0^\infty R^2(1h)R(2g)R(1i)r^2dr.$$

Thus t is negative if I_1, I_2 and I_3 all have the same sign. These integrals all do have the same sign when computed with wave functions belonging to either a square well of infinite depth or to an harmonic oscillator. It is unlikely, therefore, that the sign of G_S/G_T is negative, even if there is an appreciable admixture of ψ_b in the ground state of Bi^{210}.

The evidence against there being an appreciable amount of ψ_b in this ground state is the nonexistence of a transition to it from Pb^{210}. Though the transition to the ground state has a forty-fold energy advantage over the transition to the 47-kev level, the former has not been observed; the upper limit on its intensity is about a tenth of the intensity of the latter. How can one explain why the transition to the excited state is at least 400 times as strong as the transition to the ground state? Some of the difference is certainly caused by the selection rule which permits $\int\boldsymbol{\sigma}\cdot\mathbf{r}$ to be operative in the former and not in the latter, but the effect of the other matrix elements is of roughly the same magnitude, unless there is a cancellation of the large terms in the correction factor. According to our analysis of the spectrum of Bi^{210}, this cancellation takes place for some value of x greater than 0.35. If the ground state of Pb^{210} is pure $(\ ;g_{9/2}^2)_0$, then only the component ψ_b in the wave function of Bi^{210} has nonzero matrix elements, and x, therefore, is less than 0.1. To produce the value of x corresponding to cancellation one must form the wave function of Pb^{210} of just the right proportions of $(\ ;i_{11/2}^2)_0$ and $(\ ;g_{9/2}^2)_0$. Such a coincidence is, we believe, unlikely. Furthermore, it is not certain that the low ft-value for the decay of Pb^{210} to Bi^{210*} is consistent with an appreciable amount of $(\ ;i_{11/2}^2)_0$ in Pb^{210}.

A possible test of the amount of ψ_b in the ground state of Bi^{210} is to be found in the measurement of the half-life of the 47-kev state of Bi^{210}. From a comparison of the calculated[28] and measured[29] ratios of the numbers of $L_{\rm I}$, $L_{\rm II}$ and $L_{\rm III}$ conversion electrons, it follows that the 47-kev γ radiation is undoubtedly $M1$. The usual selection rules permit $M1$ transitions only between states of the same configuration. The half-life of the 47-kev state will be, then, inversely proportional to b^2. The Weisskopf[11] half-life for a 47-kev $M1$ transition, corrected for an internal conversion coefficient of about 20 and amended to include a better estimate of the matrix element, is 0.3×10^{-11} sec. If the nonappearance of the Pb^{210} to Bi^{210} ground state transition is taken to mean that $b^2\leq1/400$, then the half-life of the 47-kev state $\geq1.2\times10^{-9}$ sec, which is perhaps just measurable. A shorter half-life will not necessarily mean that b^2 is greater than 1/400, however, because the transition may be caused by the sort of process suggested by Austern and Sachs.[30]

I should like to thank Dr. J. M. Kennedy for valuable aid and advice, and Miss B. J. Sears for preparation of the figures and for computational assistance.

APPENDIX

A comparison of the matrix elements for the transition

$$;\ (l_1j_1)^2J_1\rightarrow(l_2j_2;l_1j_1)J_2+\beta^- \tag{4}$$

with those for the transition

$$l_1j_1\rightarrow l_2j_2+\beta^- \tag{5}$$

is desired. As is customary we shall use the symbol $|\int A|^2$ to represent the sum over all possible final states of the square of the matrix element of the operator A. In general A will be composed of a sum of terms each one of which depends on the coordinates of only one particle of the system. Each term is the product of an operator on the isotopic spin wave function and of a tensor operator of order L. The wave function of the initial state of (4) is constructed precisely as was done in reference 21. For the final state, however, we do not choose an eigenfunction of the isotopic spin operator. Instead we taken that combination of isotopic spin eigenfunctions corresponding to isotopic spins 0 and 1 which ensures that the particle in the state l_1j_1 is always a proton and that in the state l_2j_2 is always a neutron. Such a wave function, rather than any eigenfunction of isotopic spin, we believe to be appropriate to a proton-neutron system in a nucleus in the neighborhood of Pb^{208}.

Using the methods of Racah,[10] it is not difficult to show that

$$\left|\int A(2)\right|^2=2(2j_1+1)(2J_2+1)\times W^2(j_2J_2j_1J_1;j_1L)\left|\int A(1)\right|^2. \tag{6}$$

The arguments 1 and 2 refer to the one-particle and to the two-particle systems respectively. For the special case $J_1=0$, (6) becomes

$$\left|\int A(2)\right|^2=2\left|\int A(1)\right|^2\delta_{J_2L}. \tag{7}$$

Formulas (6) and (7) differ from the similar formulas of Brysk[31] and of Rose and Osborn[21] in the presence of a 2 as a factor on the right-hand side. This difference has its origin in the different definitions of the final state.

[28] Gellman Griffith, and Stanley, Phys. Rev. 85, 944 (1952).
[29] L. Cranberg, Phys. Rev. 77, 155 (1950).
[30] N. Austern and R. G. Sachs, Phys. Rev. 81, 710 (1951).
[31] H. Brysk, Phys. Rev. 90, 365 (1953).

Charged Photomesons from Carbon*

DAVID LUCKEY
Newman Laboratory of Nuclear Studies, Cornell University, Ithaca, New York
(Received September 15, 1954)

Cross sections per Mev meson energy per equivalent quantum for the production of π^+ mesons from carbon by 310-Mev bremsstrahlen have been measured in the angular range from 30° to 180°. Carbon data exist for those meson energies and angles which correspond to π^+ production from free protons by γ rays of 200, 235, and 265 Mev. A comparison is made to production from hydrogen by giving the efficiency ϵ, defined as the ratio of production per proton in carbon to that in hydrogen. Various π^-/π^+ ratios are given for carbon.

THE following measurements on the production of charged π mesons from carbon by 310-Mev bremsstrahlen are a byproduct of an experiment which measured the π^+-meson production from hydrogen using a carbon-polyethane subtraction method.[1] The measurements on carbon do not constitute a complete experiment in the sense that the complete energy spectrum of mesons at a given angle is not measured. However, the measurements are interesting since their absolute value is well determined, a wide range of angles is covered, and "complementary" hydrogen data exist at each point. These measurements supplement previous measurements on photoproduction from carbon.[2-4]

Table I shows the results for π^+-meson production from carbon. The results are grouped according to the gamma-ray energy which would produce mesons of the given energy at the given angle from free protons. These gamma-ray energies are given in the footnotes to Table I. The differential cross section is given per proton in the carbon nucleus (i.e., one-sixth the production per nucleus). The errors quoted on the differential cross section are standard deviations and contain estimates of the errors in solid angle measurement and various corrections as well as the statistical errors. The forward angles have large errors because of the uncertainty in the corrections for scattered electrons. Table I also shows the efficiency ϵ, which is the ratio of photomeson production per proton in carbon to that in hydrogen. Since this is a ratio measurement the error is due only to statistics and uncertainties in the correction for electrons. Table II shows miscellaneous π^-/π^+ ratios obtained. The π^--meson measurements were made with the same apparatus during the same run the corresponding π^+ measurements were made, only the sign of the magnetic field being changed.

These results are in agreement with those of previous experimenters where comparison can be made. The efficiency, ϵ, should be roughly independent of the details of a meson theory and depend mainly upon nuclear effects. Only part of the momentum distribution of the protons in the carbon nucleus can contribute because of energetic resons and because of the exclusion of some final states as a result of the Pauli principle. This accounts for the large decrease of the efficiency at forward angles. More detailed discussion of this point

TABLE I. π^+ photoproduction from carbon by 310-Mev bremsstrahlen. ϵ is the ratio of π^+ production per proton in the carbon nucleus to that in hydrogen at the given meson energy and angle.[a]

θ_{lab}	T(Mev)	$\frac{d^2\sigma}{d\Omega dT}\times 10^{32}$ cm² / sterad-Mev-proton-equivalent quantum	ϵ	
30°	55	2.1 ±1.4	0.34±0.25	
45°	50	2.60±0.36	0.30±0.06	
60°	45	3.30±0.32	0.46±0.06	(a)
90°	33	4.28±0.43	0.65±0.08	
120°	26	4.47±0.39	0.59±0.08	
30°	84	0.9 ±0.4	0.12±0.06	
45°	78	2.23±0.22	0.28±0.04	
60°	69	2.96±0.23	0.34±0.02	
90°	54	3.80±0.30	0.35±0.02	(b)
120°	43	3.70±0.30	0.41±0.03	
150°	37	3.65±0.30	0.40±0.05	
180°	34	3.78±0.30	0.57±0.04	
30°	114	0.1 ±0.7	0.02±0.12	
45°	107	1.62±0.26	0.20±0.05	
60°	93	2.16±0.34	0.19±0.02	
90°	70	3.42±0.29	0.28±0.02	(c)
120°	59	3.32±0.27	0.29±0.02	
150°	48	3.47±0.35	0.40±0.05	
180°	46	3.18±0.29	0.42±0.03	
90°	84	2.34±0.19	0.30±0.02	(d)

[a] π mesons at these energies and angles correspond to photoproduction from hydrogen by: (a) 200-Mev γ rays; (b) 235-Mev γ rays; (c) 265-Mev γ rays; (d) 280-Mev γ rays.

TABLE II. Various π^-/π^+ ratios from carbon with 300-Mev bremsstrahlen.

θ_{lab}	T(Mev)	π^-/π^+
30°	54	1.45±0.29
90°	34	1.22±0.07
90°	70	1.18±0.08
180°	44	1.26±0.10
180°	62	1.18±0.13

* Assisted by the joint program of the U. S. Office of Naval Research and the U. S. Atomic Energy Commission.

[1] Jenkins, Luckey, Palfrey, and Wilson, Phys. Rev. 95, 179 (1954).
[2] Peterson, Gilbert, and White, Phys. Rev. 81, 1003 (1951).
[3] J. Steinberger and A. S. Bishop, Phys. Rev. 86, 171 (1952).
[4] Feld, Frisch, Lebow, Osborne, and Clarke, Phys. Rev. 85, 680 (1951).

and other details of nucleus production can be found in work by Steinberger and Bishop,[3] Lax and Feshbach,[5] Marshak,[6] and Littauer and Walker.[7] The π^-/π^+ ratios measured here do not indicate any variation with energy or angle. However, the result of Littauer and Walker[7] at 135° and 65-Mev meson energy (1.06 ± 0.02) indicates that such a variation exists. Other experimental π^-/π^+ ratios can be found in work by Peterson *et al.*,[2] Feld *et al.*,[4] and Carothers.[8]

The author wishes to thank Thomas Jenkins, Thomas Palfrey, and Professor Robert R. Wilson for their contributions to these results.

[5] M. Lax and H. Feshbach, Phys. Rev. **81**, 189 (1951).
[6] R. E. Marshak, *Meson Physics* (McGraw-Hill Book Company, Inc., New York, 1952), Chap. 3.
[7] R. M. Littauer and D. Walker, Phys. Rev. **86**, 838 (1952).
[8] J. Carothers, Phys. Rev. **92**, 538 (1953).

PHYSICAL REVIEW VOLUME 97, NUMBER 2 JANUARY 15, 1955

Synchrotron Oscillations Induced by Radiation Fluctuations*

MATTHEW SANDS
California Institute of Technology, Pasadena, California
(Received August 23, 1954)

Phase oscillations of electrons in a high-energy synchrotron are induced by the radiation of quanta. These induced oscillations set a limit to the damping of electron bunches. This limiting bunch size is sufficient to influence the radial aperture and the radio-frequency voltage required at low beam intensities, and to reduce energy loss by coherent radiation at high intensities.

I. INTRODUCTION

THE radiation of electromagnetic energy by high-energy electrons in a synchrotron[1–5] not only requires that the radio-frequency system supply energy much in excess of that required for acceleration, but also causes a damping of the phase oscillations in normal guide fields.[2,3,5] According to the results of Bohm and Foldy,[5] in a synchrotron with a guide field which increases linearly with time, phase oscillations present at an energy E_0 will have their amplitude reduced by the factor $\exp[-(E^4-E_0^4)/E_d^4]$ at an energy E. The energy E_d may be several hundred Mev for a 1.5-Bev synchrotron with typical design parameters. Such strong damping above E_d may lead to electron bunches so small that energy loss due to coherent radiation[4,6,7] would set a serious limit to the maximum attainable beam intensity.

We present here a calculation which shows that the radiation damping will not decrease the spread in phase of an electron group below a certain minimum. This minimum is determined by the statistical fluctuations in the radiated energy loss—because of the emission of quanta. The effect is, fortunately, sufficiently large to preclude, in general, serious difficulty from coherent radiation.

We may use the following qualitative arguments to estimate the magnitude of the energy fluctuations produced by photon emission from the circulating electrons. The square of the energy fluctuation ΔE is approximately equal to the product of the mean square quantum energy ϵ_γ^2 of the radiation and the average number u of quanta emitted in one damping-time interval of the energy (phase) oscillations. This average number is ratio of the total energy radiated in one damping period to the average quantum energy. Using the relations of reference 5, we find

$$u \approx f(E/\epsilon_\gamma),$$

where E is the electron energy and f is a number which depends on the properties of the guide field and will in general be ≈ 1. We have then that

$$(\Delta E/E)^2 \approx f(\epsilon_\gamma/E).$$

From reference 13 the typical quantum energy may be taken as

$$\epsilon_\gamma \approx (\hbar c/r)(E/mc^2)^3,$$

where r is the orbit radius. We have then that

$$(\Delta E/E)^2 \approx f(\hbar/mc)(1/r)(E/mc^2)^2.$$

The oscillations in radius and phase are simply related to the energy fluctuations.

We present below more detailed calculations, including also the effects of harmonic accelerating fields and noncircular guide fields. We consider specifically only synchrotrons with constant-gradient ("weak-focusing") guide fields.

* This work was supported in part by the U. S. Atomic Energy Commission.
[1] D. Iwanenko and I. Pomeranchuk, Phys. Rev. **65**, 343 (1944).
[2] J. Schwinger, Phys. Rev. **70**, 798 (1946).
[3] N. Frank, Phys. Rev. **70**, 177 (1946).
[4] L. Schiff, Rev. Sci. Instr. **17**, 6 (1946).
[5] D. Bohm and L. Foldy, Phys. Rev. **70**, 249 (1946).
[6] J. Nodvick and D. Saxon, University of California at Los Angeles Technical Report No. 21, May 1954; Phys. Rev. **96**, 180 (1954).
[7] J. Schwinger (1945, unpublished).

II. EQUATIONS OF MOTION

We consider a magnetic guide field composed of circular sectors connected by field-free regions of total length L. In the circular sectors the synchronous electrons move with a radius r_s and an angular velocity ω_s determined by the (angular) frequency of the accelerating system $\omega_0=k\omega_s/\lambda$, where k is an integer (the harmonic number), and $\lambda=1+L/2\pi r_s$. We take further that the guide field at the synchronous orbit increases linearly with time and that the field varies near r_s as $(r/r_s)^{-n}$.

Following Blachman and Courant,[8] we may write the equations which determine the phase oscillations in terms of small deviations from synchronous values as

$$\Delta E/E_s=(1-n)(\Delta r/r_s); \tag{1}$$

$$\Delta r/r_s=-\Delta\omega/\omega_s=-(\lambda^2/k)(1/\omega_s)(d\phi/dt), \tag{2}$$

where ϕ is the phase angle of the electron with respect to the accelerating field, and E is the energy of the electron which is assumed to be much greater than mc^2.

Differentiating and combining Eqs. (1), (2), we have

$$\frac{d^2\phi}{dt^2}+\frac{k\omega_s}{(1-n)\lambda^2E_s}\frac{d}{dt}(\Delta E)+\frac{\dot{E}_s}{E_s}\frac{d\phi}{dt}=0, \tag{3}$$

where $\dot{E}_s$ is the constant rate of increase of the synchronous energy.

The rate of change of the energy of the electron is given by the power received from the radio-frequency system and from the variation of flux linking the orbit, less the power lost by radiation.

$$(d/dt)\Delta E=P_a(\phi,d\phi/dt)-P_\gamma(d\phi/dt,t)-\dot{E}_s, \tag{4}$$

where the t-dependence in P_γ represents the variations —due to quantum emission—in the rate of energy loss. As the period of a phase oscillation is long compared to a revolution period, we follow the usual procedure and take for dE/dt its average over one revolution.

The power received by the electron is then (correct to first order terms)[9]

$$P_a(\phi,d\phi/dt)=P_{as}+P_{as}\Delta\phi\cot\phi_s+(P_{as}-\dot{E}_s)(\lambda/k\omega_s)(d\phi/dt), \tag{5}$$

where ϕ_s is the phase of the synchronous electron and P_{as} is the power received by it from the radio-frequency system. We have for convenience taken the flux linking the equilibrium orbit to be zero. Equation (9), below, would remain unchanged in any case.

We represent the power radiated by the electron as the sum of its average value over one revolution, which varies only slowly with time, and a remaining term $p(t)$ which contains the rapid variations due to quantum emission. We ignore the dependence of $p(t)$ on $d\phi/dt$ using instead the function pertinent to the synchronous electron:

$$P_\gamma(d\phi/dt,t)=\langle P_\gamma(d\phi/dt)\rangle_{Av}+p(t). \tag{6}$$

The justification—and necessity—for such a procedure in problems of this type is given by Chandrasekhar.[10]

The average power[1,2] is given by

$$\left\langle P_\gamma\left(\frac{d\phi}{dt}\right)\right\rangle_{Av}=\frac{4\pi e^2}{3}\frac{1}{r}\left(\frac{E}{mc^2}\right)^4\frac{c}{2\pi r+L}, \tag{7}$$

where c is the velocity of light. We may expand this for energies near E_s and, using (1) and (2), obtain

$$\left\langle P_\gamma\left(\frac{d\phi}{dt}\right)\right\rangle_{Av}=P_{\gamma s}\left[1-\left(3-4n-\frac{1}{\lambda}\right)\frac{\lambda^2}{k\omega_s}\frac{d\phi}{dt}\right], \tag{8}$$

where $P_{\gamma s}$ is the average power loss (at the energy E_s) by the synchronous particle.

Combining Eqs. (3), (4), (5), (6), and (8), we have the differential equation of the phase oscillations.

$$\frac{d^2\phi}{dt^2}+[(3-4n)P_{\gamma s}+(1-n)\dot{E}_s]\frac{1}{(1-n)E_s}\frac{d\phi}{dt}+\frac{k}{\lambda^2}\frac{\omega_s}{(1-n)E_s}[P_{as}\cot\phi_s\Delta\phi-p(t)]=0. \tag{9}$$

In Eq. (9) the coefficients of $d\phi/dt$ and $\Delta\phi$ depend on E_s which varies slowly with time. We may treat the coefficients as constant provided that the change in E_s is small in one damping time constant (i.e., a time equal to the reciprocal of the coefficient of $d\phi/dt$). Since $P_{\gamma s}$ increases as E_s^4 such a treatment will be valid for energies appreciably above a critical energy E_c defined by

$$P_{\gamma s}(E_c)=(1-n)\dot{E}_s/(3-4n). \tag{10}$$

For energies below E_c the equation may be treated by the methods of references 5 or 8, with which one obtains the usual adiabatic damping of the phase oscillations.

Restricting ourselves to energies above E_c we may neglect in Eq. (9) the terms in $\dot{E}_s$. We have then

$$d^2\psi/dt^2+\rho d\psi/dt+\Omega^2\psi=g(t), \tag{11}$$

where

$$\psi=\Delta\phi=\phi-\phi_s,$$

$$\rho=\frac{3-4n}{1-n}\frac{P_{\gamma s}}{E_s},$$

$$\Omega^2=aP_{as}\cot\phi_s,$$

$$g(t)=ap(t),$$

$$a=\frac{k}{\lambda^2}\frac{\omega_s}{(1-n)E_s}.$$

[8] N. Blachman and E. Courant, Rev. Sci. Instr. **20**, 596 (1949).

[9] L. J. Laslett kindly pointed out that in the original manuscript the term in $\dot{E}_s$ had been omitted from this equation.

[10] S. Chandrasekhar, Revs. Modern Phys. **15**, 1 (1943).

III. FLUCTUATION EFFECTS

We now ignore any oscillations in phase which may be present initially (and which would be rapidly damped) and determine the phase oscillations which arise due to the fluctuation in the energy loss as represented by $g(t)$. We compute only the equilibrium value of these oscillations. It may be expected that they will reach this equilibrium value in a time about $1/\rho$.

The driving function $g(t)$ may be expressed as an infinite sum, each term of which represents the emission of a quantum of energy ϵ_i assumed to be emitted instantaneously.

$$g(t)=a\sum \epsilon_i\delta(t-t_i)-\bar{g}.$$

The resulting phase oscillations will be given by

$$\psi(t)=a\sum \epsilon_i H(t-t_i)-\phi_s,$$

where $H(t)$ is the response to a unit impulse of the system described by Eq. (11).

Let $u(\epsilon)d\epsilon$ represent the number of quanta per unit of time of energy between ϵ and $\epsilon+d\epsilon$. Applying Campbell's theorem,[11] the contribution of these quanta to the mean square fluctuation in phase is

$$d\langle\psi^2\rangle_{Av}=u(\epsilon)d\epsilon\int_{-\infty}^{+\infty}a^2\epsilon^2H^2(t)dt, \tag{12}$$

if it is assumed that the quanta are emitted statistically independently.[12] The total fluctuation in phase is

$$\langle\psi^2\rangle_{Av}=a^2\int_0^\infty\epsilon^2u(\epsilon)d\epsilon\int_{-\infty}^{+\infty}H^2(t)dt. \tag{13}$$

The integral over t is conveniently expressed as

$$\frac{1}{2\pi}\int_{-\infty}^{+\infty}\frac{d\sigma}{|\Omega^2-\sigma^2+i\rho\sigma|^2}=\frac{1}{2\Omega^2\rho}. \tag{14}$$

The quantum distribution function $u(\epsilon)$ is related to the power spectrum $p(\epsilon)$ of the radiation by $\epsilon u(\epsilon)=p(\epsilon)$. Taking the power spectrum given by Schwinger,[13]

$$u(\epsilon)=\frac{3^{5/2}}{8\pi}\frac{P_{\gamma s}}{\epsilon_c^2}\int_{\epsilon/\epsilon_c}^\infty K_{5/3}(\eta)d\eta dx. \tag{15}$$

with

$$\epsilon_c=(3/2)\hbar\omega_s(E_s/mc^2)^3. \tag{16}$$

The integral over ϵ in Eq. (13) is then

$$(3^{5/2}/8\pi)P_{\gamma s}\epsilon_c\int_0^\infty x^2\int_x^\infty K_{5/3}(\eta)d\eta dx. \tag{17}$$

Evaluating the definite integral,[14] we obtain

$$\int_0^\infty\epsilon^2u(\epsilon)d\epsilon=(55/2^33^{\frac{1}{2}})P_{\gamma s}\epsilon_c. \tag{18}$$

Combining the results of (13), (14), and (18), and recalling the definitions of (11), we obtain finally:

$$\langle\psi^2\rangle_{Av}=\frac{55\sqrt{3}}{2^6}\frac{\hbar c}{e^2}\frac{k}{\lambda}\frac{\tan\phi_s}{(3-4n)}\frac{mc^2}{E_s}. \tag{19}$$

where, in keeping with our approximations, we have set $P_{as}=P_{\gamma s}$. We shall use this result later to compute the energy loss due to coherent radiation by the electron bunch.

It will be recalled that (19) is applicable for energies greater than E_c defined by (10). We may obtain an explicit expression for E_c, using [from (7)]

$$P_{\gamma s}=\frac{2}{3}\frac{c}{\lambda}\frac{e^2}{r_s^2}\left(\frac{E_s}{mc^2}\right)^4, \tag{20}$$

whence

$$\frac{E_c}{mc^2}=\left[\frac{3}{2}\frac{\lambda}{c}\frac{r_s^2}{e^2}\frac{(1-n)}{(3-4n)}\dot{E}_s\right]^{\frac{1}{4}}. \tag{21}$$

It is perhaps more convenient to express (21) in terms of the net energy gain per turn,

$$\Delta E_0=\lambda(2\pi r_s/c)\dot{E}_s.$$

Then

$$\frac{E_c}{mc^2}=\left[\frac{3}{4\pi}\frac{(1-n)}{(3-4n)}\frac{r_s}{r_0}\frac{\Delta E_0}{mc^2}\right]^{\frac{1}{4}}. \tag{22}$$

where r_0 is the classical electron radius. It may be expected that for energies above E_c the relative error in (19) will be of the order of $\Delta E_0/\Delta E_\gamma$, where ΔE_γ is the energy radiated in one turn.

The oscillations in radius associated with the fluctuations in phase may also be of some interest. From (2) we have

$$\left\langle\frac{(\Delta r)^2}{r_s^2}\right\rangle_{Av}=\frac{\lambda^4}{k^2}\frac{1}{\omega_s^2}\left\langle\left(\frac{d\psi}{dt}\right)^2\right\rangle_{Av}.$$

The damping of the phase oscillations in one cycle is very small.[15] We may, therefore, write

$$\langle(d\psi/dt)^2\rangle_{Av}=\Omega^2\langle\psi^2\rangle_{Av}.$$

We have then, for the mean square fluctuation in radius,

$$\left\langle\frac{(\Delta r)^2}{r_s^2}\right\rangle_{Av}=\frac{55}{2^5\sqrt{3}}\frac{\hbar c}{e^2}\frac{1}{(1-n)(3-4n)}\frac{r_0}{r_s}\left(\frac{E_s}{mc^2}\right)^2. \tag{23}$$

[11] See, e.g., S. O. Rice, Bell System Tech. J. **23**, 282 (1944).
[12] See appendix for further discussion of this point.
[13] J. Schwinger, Phys. Rev. **75**, 1912 (1949), Eqs. (11), (16).
[14] G. N. Watson, *Bessel Functions* (McMillan Company, New York, 1945), p. 388.
[15] The logarithmic decrement is $\approx(\Delta E_\gamma/E_s)^{\frac{1}{2}}$. See, e.g., reference 5.

IV. NUMERICAL EXAMPLES

Design parameters for a 1.5-Bev synchrotron might be

$$r_s=375 \text{ cm}, \quad k=4, \quad \lambda=1.25,$$
$$\Delta E_0=750 \text{ volts}, \quad n=0.6.$$

For these values the fluctuations will become significant above $E_c=380$ Mev [using (22)] and will become dominant by about $E_s=700$ Mev.

The mean square fluctuation in phase will then be given by

$$\langle\psi^2\rangle_{Av}=1080(mc^2/E_s)\ \tan\phi_s=(0.56 \text{ Bev}/E_s)\ \tan\phi_s.$$

For 1.5 Bev the root-mean-square angle is

$$\langle\psi^2\rangle_{Av}^{\frac{1}{2}}=0.59(\tan\phi_s)^{\frac{1}{2}}. \tag{24}$$

Thus, an equilibrium phase angle of about $\pi/6$ is required if electrons are not to be lost due to these fluctuations. The maximum voltage of the radio-frequency system must, therefore, exceed 200 kilovolts (since $\Delta E_\gamma=120$ kev for $E_s=1.5$ Bev).

The required radial aperture at high energies may be determined from (23). For the parameters chosen here,

$$\left\langle\frac{(\Delta r)^2}{r_s^2}\right\rangle_{Av}=4.2\times10^{-13}\left(\frac{E_s}{mc^2}\right)^2=10^{-4}\left(\frac{E_s}{8 \text{ Bev}}\right)^2.$$

For $E_s=1.5$ Bev the root-mean-square spread in radii is

$$\langle(\Delta r)^2\rangle_{Av}^{\frac{1}{2}}=1.9\times10^{-3}r_s=0.7 \text{ cm}. \tag{25}$$

The effective magnetic aperture must be many (≈10) times this value if loss of electrons is to be avoided.

V. COHERENT RADIATION

Several authors[4,6,7] have considered the increase in the rate of energy loss, above that given by P_γ, due to the coherent effects of all the electrons in a bunch. Schiff[4,6] gives expressions for the case in which the electrons are spread in a Gaussian distribution about the equilibrium phase. Since the important contribution to the coherent radiation comes from long wavelengths, the results depend on the proximity of metallic surfaces. If we assume that the mean square fluctuation in the phase of a single electron represents the longitudinal spread of a group of electrons, then the additional loss in one revolution *for each electron* (due to coherent radiation) is, in the absence of metallic shielding,[4]

$$\Delta E_{coh}=3^{1/6}\Gamma^2(2/3)(r_0/r_s)k^{\frac{1}{2}}N\langle\psi^2\rangle_{Av}^{-\frac{1}{4}}mc^2,$$

where N is the total number of electrons accelerated. Using the results above [Eq. (24)] and taking $\phi_s=\pi/4$, we see that the coherent loss per electron is, at 1.5 Bev,

$$\Delta E_{coh}=2.5\times10^{-9}N \text{ electron volts}.$$

Typical intensities involve the acceleration of from 10^8 to 10^{11} electrons. Even for $N=10^{12}$, the coherent energy loss is only 2500 ev, which is negligible in comparison with the incoherent losses.[16]

If the electron orbits are within a vacuum chamber which has conducting walls whose separation is less than $(r_s/k)\langle\psi^2\rangle_{Av}^{\frac{1}{2}}$ (as would normally be) the coherent loss is less than the above value. In a typical case the shielded coherent loss is computed (following reference 6) to be less than 1/50 the nonshielded loss.

This work profitted from discussions with M. Bloch, Leverett Davis, Jr., R. P. Feynman, R. V. Langmuir, Using the results above [(24)] and taking $\phi_s=\pi/4$, Robert L. Walker, and Kenneth M. Watson. L. Jackson Laslett kindly pointed out the existence of some numerical errors in the original manuscript.

APPENDIX[17]

The justification for the addition of quantum recoil effects to a "classical" system is contained essentially in the work of Bloch and Nordsieck and others.[18] These authors show that emissions of individual quanta from a classical radiating system are statistically independent.

Since we are here interested in the fluctuations in energy loss in a time interval much longer than time of emission of a typical quantum, the uncertainty in the definition of the electron energy is insignificant compared with the calculated energy fluctuations.

Watson[19] points out that a complete quantum-mechanical calculation can be performed by introducing into the equation of the phase oscillation the interaction of the electron with the radiation field. This results in replacing $g(t)$ in Eq. (11) by $a(e\mathbf{v}\cdot\boldsymbol{\mathcal{E}}-e\mathbf{v}\cdot\langle\boldsymbol{\mathcal{E}}\rangle_{Av})$, where $\mathbf{v}$ is the electron velocity, $\boldsymbol{\mathcal{E}}$ is the radiation field strength, and $\langle\boldsymbol{\mathcal{E}}\rangle_{Av}$ is its expectation value. An evaluation of the quantum-mechanical expectation value of ψ^2 then yields the result obtained here for the time average.

---

[16] For intensities as large as 10^{12} electrons, spreading of the bunch due to space-charge forces may begin to become significant.

[17] I am indebted to K. M. Watson for the illuminating discussions which led to the following remarks.

[18] F. Bloch and A. Nordsieck, Phys. Rev. **52**, 54 (1937); W. Pauli and M. Fierz, Nuovo cimento **15**, 167 (1938); R. Glauber, Phys. Rev. **84**, 395 (1951).

[19] Kenneth M. Watson (private communication).

Characteristics of Bubble Chambers*

DONALD A. GLASER AND DAVID C. RAHM†
Randall Laboratory of Physics, University of Michigan, Ann Arbor, Michigan
(Received September 23, 1954)

The bubble chamber is a new radiation detector in which ionizing events produce tracks consisting of strings of tiny bubbles in a superheated liquid. By means of fast flash photography, practically distortionless bubble tracks can be recorded for the study of high-energy nuclear events. A chamber six inches long has been constructed and chambers several feet long seem feasible, so the bubble chamber can have both high stopping power and large size. Liquids containing only elements of low atomic number can be used to minimize Coulomb scattering so that accurate magnetic curvature measurements can be made. Tracks of minimum ionizing particles contain up to 100 bubbles per centimeter, and the bubble density varies with ionization density, thus making bubble counting a possible method for measuring ionization. The short sensitive and resetting times of bubble chambers tend to reduce background radiation problems and increase data collection rates for experiments with pulsed accelerators. Unfortunately the lifetime of the bubble nuclei seems to be too short to permit counter-controlled expansion after the passage of the particle to produce tracks. Therefore it is unlikely that the types of bubble chambers described here will be useful for cosmic-ray experiments.

INTRODUCTION

A BUBBLE chamber is a vessel filled with a transparent liquid which is so highly superheated that an ionizing particle moving through it starts violent boiling by initiating the growth of a string of bubbles along its path. For the chambers to be described below, these bubble tracks provide an accurate record of the paths of ionizing particles when the bubbles are photographed a few microseconds after the ionizing events. Although still in its infancy, the bubble chamber technique shows considerable promise of being useful for experiments in high-energy nuclear physics because it combines many of the advantages of cloud chambers and nuclear emulsions with several new unique characteristics.

Bubble chambers used in the early studies contained only a few cubic centimeters of liquid,[1] but recently a chamber with a sensitive volume of over 600 cubic centimeters has been operated successfully, and chambers up to a few feet long seem feasible. Thus liquid-filled bubble chambers combine the high stopping power of nuclear emulsions with the large volume of cloud chambers. The density of typical operating liquids is about 0.5 g/cc, which is 300 times the density of the gas in the usual argon-filled cloud chamber. When the operating liquid is sufficiently superheated to form tracks of minimum ionizing particles, the bubbles grow so rapidly that a photograph can be taken before there is time for motions in the liquid to distort the tracks. This fact, coupled with the possibility of using liquids containing only elements of low atomic number to minimize Coulomb scattering, should permit very accurate measurements of magnetic curvature. On the other hand, liquids containing various atomic nuclei can be used as the sensitive media for special experiments. For example, beautiful bubble tracks have been obtained in a chamber filled with liquid hydrogen,[2] a pure proton target.

Since the sensitive time of these chambers is usually less than 10 milliseconds, and they can be reset in less than 5 seconds, they are ideally suited for experiments with pulsed accelerators in which the interesting events occur in short bursts in a sea of rather intense background radiation.

Bubble tracks made by many kinds of particles have been photographed, including minimum ionizing electrons and muons, and slow nuclear fragments. The density of bubbles along a track is found to increase with ionization density, but no careful calibration has been made. Particles originating in radioactive materials, pulsed accelerator beams, and cosmic rays have been photographed in bubble chambers, but counter-controlled expansions after the passage of a particle have not produced visible tracks.

GENERAL PRINCIPLES OF BUBBLE CHAMBER OPERATION

Track formation in a bubble chamber can be understood in terms of two fundamental facts; first, a clean liquid can be superheated considerably by rapidly reducing the pressure on it, and second, an ionizing event can initiate boiling in a liquid if it is sufficiently superheated. The discovery of this last effect[3] was motivated by the desire to produce a detector with properties like those of the bubble chamber, and was guided by a detailed physical model of the mechanism by which ionization could nucleate bubble formation.[4] Since the first experiments with diethyl ether, other liquids including iso-pentane, benzene, sulfur dioxide, and ethyl alcohol have been found to be similarly radiation-sensitive, in close agreement with the pre-

* This research has been supported in part by the Michigan Memorial-Pheonix Project and the National Science Foundation.
† U. S. Atomic Energy Commission Predoctoral Fellow 1951–1953.
[1] D. A. Glaser, Phys. Rev. 91, 762 (1953).
[2] J. G. Wood, Phys. Rev. 94, 731 (1954).
[3] D. A. Glaser, Phys. Rev. 87, 665 (1952).
[4] D. A. Glaser, Nuovo cimento 11, Suppl 2, 361 (1954).

dictions of the theory. Even liquid hydrogen and liquid nitrogen have been found to be radiation-sensitive[6] in the expected range, although temperatures and pressures were not measured accurately in these tests. To find out whether the proposed explanation of the effect is correct, a number of liquids are being exposed to various radiations in an extended series of experiments now in progress. The results thus far are in good agreement with the theory, though other plausible models cannot yet be excluded by the experiments. One striking result is that diethyl ether can be superheated only slightly when exposed to fast neutrons. A higher degree of superheat is attainable in the presence of a gamma-ray flux, the results being independent of the gamma-ray energy in the range 0.6 Mev to 5 Mev. Finally an ultimate limit is reached at which a liquid will not remain superheated even in the absence of external sources or radiation. This limit, which we have called the absolute superheat limit, is analogous to the "cloud limit" in a cloud chamber, and probably corresponds to the condition under which bubbles are produced copiously by spontaneous density fluctuations. The results of this series of experiments should be ready for publication soon.

RESUME OF EARLY EXPERIMENTS

1. Experimental Technique

Until recently all of our bubble chambers have been made entirely of pyrex glass. Various sizes of heavy-walled tubing were used to make the chamber body which was joined to heavy-walled capillary tubing. The capillary was in turn sealed with O-rings to pressure-controlling apparatus operated at room temperature. After fabrication the glass parts were cleaned with hot chromic acid, flushed with distilled water, evacuated, flamed under hard vacuum, and sealed off. Just before use the seal was broken while immersed in the filling liquid. In the first experiments the pressure was controlled by a $\frac{1}{2}$-inch steel piston sealed by an O-ring into a brass cylinder and adjusted by a screw and hand crank. For automatic operation the control apparatus contained a small teflon diaphragm backed with neoprene and actuated by compressed air as in a cloud chamber. Temperatures in the range 135°C to 143°C were maintained by an oven or liquid bath for experiments with diethyl ether.

2. The First Particle Tracks

In the first attempts to observe tracks the chamber was superheated by reducing the pressure rapidly from twenty atmospheres to one atmosphere while the chamber temperature was 135°C. An ether-filled chamber with a sensitive volume of few cubic centimeters often remains quietly superheated for several seconds under these conditions. High-speed movies show that the eruption which finally occurs sometimes begins with a line of bubbles passing through the liquid.[4] The track develops very rapidly so that after twenty milliseconds it is hardly recognizable.

3. Counter-Controlled Pictures

Cosmic-ray tracks[1] were next obtained in automatic operation by means of a vertical counter telescope which was permitted to trigger flashlamps during the few seconds of sensitive time if a particle passed through the chamber. Variation of the temperature in these experiments showed that the bubble density along a minimum-ionizing track increases rapidly with temperature.

4. Vibration-Controlled Pictures

In order to see events more interesting than muons passing straight through the chamber, we took advantage of the violence of the eruption which produces an audible "plink" at each event. A General Electric variable-reluctance phonograph pickup was mounted with its stylus pressing against the wall of the chamber. Vibration signals occurring during the quiescent period after the expansion were allowed to trigger the lights and take pictures. In this way we saw tracks of particles passing through the chamber in various directions, being scattered, and so on, but the method is limited intrinsically by the speed of sound in the liquid. Vibration signals require about 100 microseconds to go a distance of 10 cm in the liquid, but in 10 microseconds the bubbles are already too big for precision measurements.

5. Scattered-Light-Controlled Pictures

Another method of triggering the flashlamps for random events in a nonselective way made use of light scattered by the bubbles. A weak steady light beam passing through the chamber was monitored by a photomultiplier which flashed the lights whenever it saw any change in optical density. This method gave rather poor quality tracks but could probably be refined.

CONSTRUCTION AND OPERATION OF LARGE BUBBLE CHAMBERS

Fabrication of large all-glass chambers with flat viewing windows is a formidable technical problem because the chambers are required to withstand rapid pressure changes of 300 pounds per square inch. On the other hand chambers built with metal parts, cements, and gaskets would not remain quiescently superheated under conditions required for the detection of minimum-ionizing particles. Boiling always started at some wall surface in such a chamber before the required degree of superheat could be attached.

For operation with diethyl ether, a solution of the problem was to expand the chamber very rapidly.

[6] R. H. Hildebrand and D. E. Nagle, Phys. Rev. 92, 517 (1953).

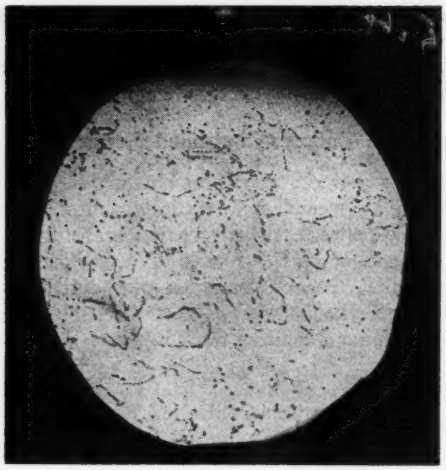


FIG. 1. Electron tracks produced in an ether-filled bubble chamber by gamma rays from a 25-mC radium source held about three inches from the center of the chamber in the plane of the photograph. The chamber has an inside diameter of two inches, a depth of one inch, and was operated at a temperature of 155°C and a pre-expansion pressure of 320 pounds per square inch (gauge). This picture was taken 5.4 milliseconds after the beginning of the rapid expansion.

Boiling then begins on the walls, but fortunately there is an interval of time during which the pressure is low enough to permit the formation of tracks of minimum-ionizing particles. The same effect has been observed with liquid hydrogen chambers.[2] The explanation for this behavior is probably to be found in the fact that the growth of a bubble in the boiling process is severely retarded after the earliest stages by heat conduction in the liquid, a relatively slow process. A recently developed mathematical theory of bubble growth incorporating the heat conduction effect has given dramatic agreement with measured bubble growth rates.[6,7]

A chamber based on these principles has been constructed with plate glass windows sealed to a duraluminum body with teflon gaskets. The sensitive volume, which is two inches in diameter and one inch thick, communicates with an expansion chamber outside the oven through a $\frac{3}{8}$-inch brass tube eight inches long. Nitrogen at a pressure of 300 pounds per square inch compresses the chamber by pushing on a flexible teflon diaphragm backed with neoprene in the expansion chamber. Rapid expansions are carried out by releasing the nitrogen with a magnetic pop valve of the type commonly used for cloud chambers. With this arrangement minimum-ionization tracks due to a radium source standing nearby can first be seen about five milliseconds after the expansion starts. Figure 1 is a photograph taken 5.4 milliseconds after the beginning of the expansion. Tracks of various ages appear in a later photograph taken at 6.7 milliseconds (Fig. 2). A krypton-filled flashlamp with a flash duration of about 5 microseconds illuminated a diffusing screen just behind the chamber for these bright field pictures. Linagraph panchromatic recording negative was used in a 35-mm camera operated at $f/16$. Photographs taken as late as 12.0 milliseconds from the beginning of the expansion still show new tracks, so the sensitive time of this chamber is about 7.0 milliseconds.

For experiments at the Brookhaven Cosmotron a bubble chamber with a sensitive volume of six by three by two inches has been built. Made of one piece of dural, the body of the chamber has walls one inch thick polished on the inside. Tracks are photographed through tempered plate glass windows of free area six by three inches bolted to the chamber body and sealed with teflon gaskets. At one end of the chamber the two by three inch inside cross section is flared out to a circular cross section with a gasket ring five inches in diameter. A diaphragm of 30-mil teflon backed by a sheet of $\frac{1}{8}$-inch silicone rubber is clamped to this ring, effectively forming a movable wall of the chamber. The position of this diaphragm is controlled by compressed nitrogen. When released by a magnetic pop valve, the nitrogen escapes and permits the diaphragm to be pushed by the hot liquid against a fixed supporting plate. The controlling magnet for the expansion valve is outside the oven, but the valve and its teflon seat are in the oven. Heat is provided by conducting glass plates clamped to the flat metal chamber walls with thin mica separators. A thermistor inserted in a well

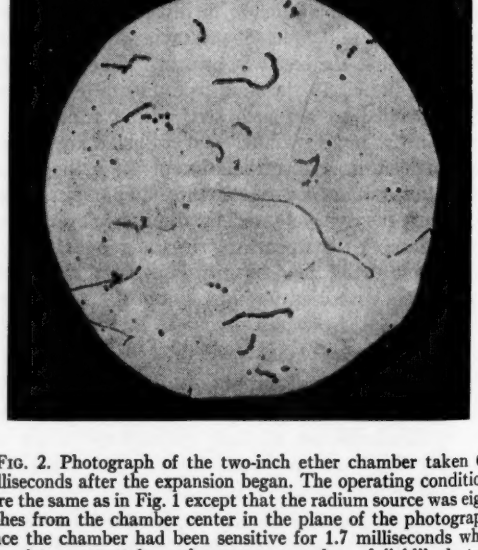


FIG. 2. Photograph of the two-inch ether chamber taken 6.7 milliseconds after the expansion began. The operating conditions were the same as in Fig. 1 except that the radium source was eight inches from the chamber center in the plane of the photograph. Since the chamber had been sensitive for 1.7 milliseconds when this picture was taken, there are a number of "old" electron tracks characterized by larger bubbles than mark the "recent" tracks.

[6] H. K. Forster and N. Zuber, J. Appl. Phys. 25, 474 (1954).
[7] M. S. Plesset and S. A. Zwick, J. Appl. Phys. 25, 493 (1954).

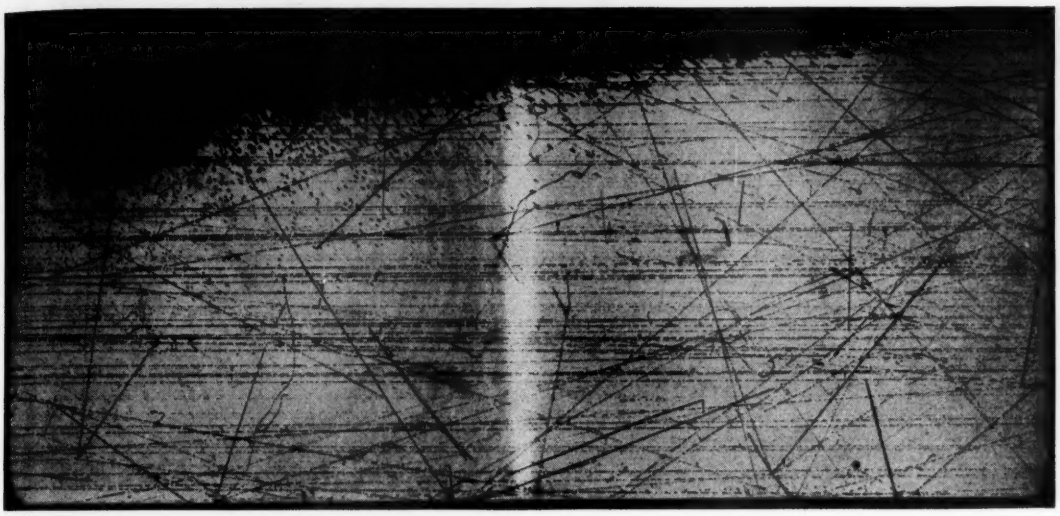


FIG. 3. A beam of 2-Bev protons from the Brookhaven Cosmotron enters an iso-pentane chamber six inches long from the left in this picture and produce a number of nuclear interactions in the liquid. Tracks of various bubble densities are produced in this chamber, operated at 157°C, even though violent boiling has already started at the top.

in the dural chamber wall controls the power delivered to these heaters by means of a conventional 60-cycle ac bridge circuit and a 5632 thyratron. Finally the whole assembly is insulated with a blanket of glass wool in a thin-walled plywood box provided with double "storm windows" to reduce heat losses from the windows of the bubble chamber. This construction, in which the moving diaphragm is in the oven, is a return to the typical expansion cloud chamber type of design which permits more rapid expansion than the earlier designs in which the liquid was expanded through a tube running to an expansion device outside the oven. When operated at 157°C with a pre-expansion pressure of 350 pounds per square inch (gauge) using iso-pentane, the six-inch chamber is fully sensitive to minimum ionizing particles 3.5 milliseconds after the expansion is initiated. The chamber seems to remain sensitive for about 10 milliseconds.

Figure 3 is a picture taken in an external beam of 2-Bev protons at the Brookhaven Cosmotron. The iso-pentane-filled chamber was expanded four milliseconds before the beam arrival and the lights were flashed 90 microseconds after a four-fold scintillation counter telescope recorded the arrival of the first hundred protons of the pulsed beam on a fast scaler. In this picture the 2-Bev protons entering at the left make tracks containing about 24 bubbles per centimeter in isopentane at 157°C while at 155°C the bubble density is closer to 8 bubbles per centimeter along such tracks. A number of nuclear interactions in the liquid can be seen to emit both lightly and heavily ionizing particles. Violent boiling began first at the top of the chamber partly because more heat was inadvertently delivered to the top wall. Figure 4 is a picture taken earlier with respect to the proton pulse, but the conditions are otherwise the same as for Fig. 3. One of the three visible particles apparently originating in an interaction in the wall of the chamber exhibits what seems to be a characteristic successive π-μ-electron decay. Eight similar π-μ-electron events were found in a rapid scanning of the first 22 pictures taken in this test at the Cosmotron. Pictures were taken every 30 seconds in this run, but more rapid operation is possible.

These pictures were taken on 70 mm linagraph orthofilm with a stereoscopic camera operated at $f/45$ using bright field illumination.

There seems to be no fundamental limitation on the possible size of bubble chambers constructed and operated in this way. A considerably larger chamber is being planned for use at the Cosmotron.

SENSITIVE TIME OF BUBBLE CHAMBERS: COUNTER-CONTROLLED EXPANSIONS

More precise information about the sensitive time of the two-inch chamber was obtained by exposing it to intense bursts of x-rays generated by the pulsed 420-kilovolt electron injector of the University of Michigan synchrotron. X-ray pulses several microseconds long were introduced at various times during the expansion process to see how the chamber sensitivity varied as the pressure was dropping. Because of limitations of photographic contrast and grain size, the lower limit of size of bubbles observed in this particular experiment was 0.1 mm, though that is not an ultimate limit for either photographic recording or bubble chambers. By counting the bubbles in sample volumes of the bubble chamber, the total number of bubbles in the chamber was estimated to be zero for

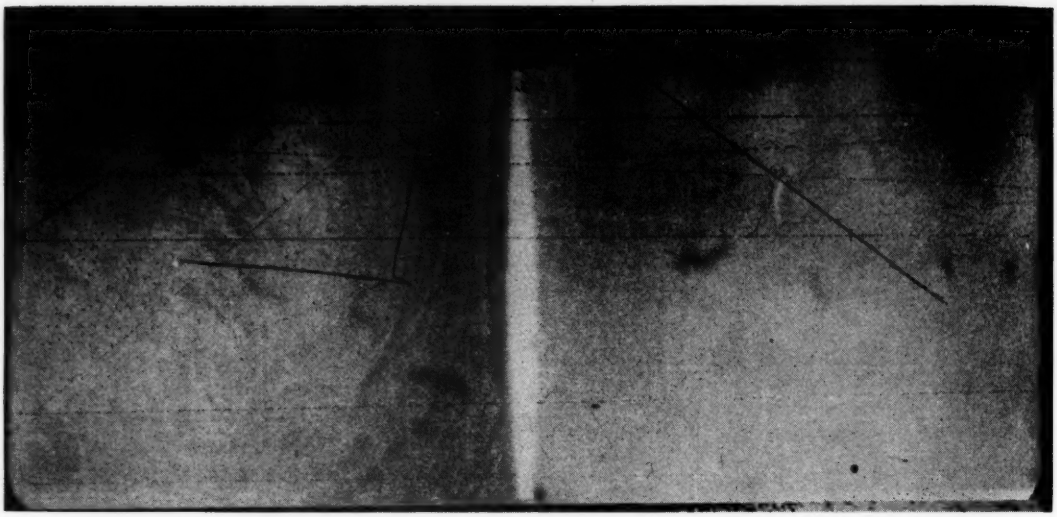


FIG. 4. A nuclear interaction in the top wall of the six-inch chamber operated in the 2-Bev proton beam from the Cosmotron produces a particle which stops in the liquid and produces an event which is probably a π-μ-e decay. The first 22 pictures of this run contained 8 examples of this type of event.

radiation bursts earlier than 3.6 milliseconds after the expansion began, 20 for bursts at 3.7 milliseconds, 20 000 at 4.9 milliseconds, and 20 000 for bursts occurring until 12.0 milliseconds at which time the chamber bubble count drops very rapidly to zero. These results are independent of the time of the light flash so long as it comes at least a few microseconds after the radiation burst and not earlier than 4.5 milliseconds after the expansion begins. The rapid cutoff in bubble sensitivity is simultaneous with a pressure rise in the chamber which is accompanied by the appearance of a mushroom-shaped jet of vapor-liquid foam at the exhaust tube orifice.

We can understand the plateau and final cutoff in terms of the pressure curve. Apparently there is some time interval during the cycle in which wall boiling evolves vapor just fast enough to keep up with the expansion process, maintaining a sufficiently low pressure to provide the superheat condition necessary for minimum ionization track formation. With the small all-glass bubble chambers, this pressure is known to be one atmosphere, is maintained for several seconds, and requires a liquid temperature of about 143°C. For rapid operation of the large metal chambers the ambient pressure during the sensitive time must be higher, and is estimated to be about 10 atmospheres when the operating temperature is 155°C. At 160°C the whole chamber exceeds the absolute superheat limit sometime during the expansion before tracks can be seen and is filled with a dense fine-grained foam regardless of the intensity of ionization in the liquid.

It is important for cosmic-ray studies with bubble chambers to be able to take counter-controlled pictures of sizeable chambers. Since the large chamber described here has a sensitive time of only a few milliseconds, counter-controlled expansions after the event of interest seem necessary if this type of chamber is to be used. Whether this will be feasible depends crucially on the lifetime in the unexpanded state of the bubble nuclei produced by the ionizing event. Presumably the visible bubbles originate from microscopic nuclei that grow to visible size only after the pressure has been lowered sufficiently. Their lifetime is limited by ion recombination times or ion mobilities if the nucleating mechanism is electrical, or by thermal diffusivity if the nucleating mechanism is thermal. It is probable that an ionizing event creates bubble nuclei containing different amounts of energy, so that the more energetic ones require less severe superheat to produce visible bubbles. Using these ideas we can explain the results of the sensitivity experiment by supposing that at each instant during the expansion process the conditions are right for rapid growth to stable size of nuclei of a certain energy. Stronger ones grow to be surviving bubbles and weaker ones disappear. Then the observed bubble density *versus* time curve is a result of selecting from a population of nuclei of various energies those which are "strong enough" to grow at the pressure and temperature prevailing at the time of their "birth." If by lifetime we mean the length of time a nucleus remains capable of forming a bubble able to grow when given a chance, then the date are consistent with a lifetime of less than a few microseconds. On the other hand, this experiment proves definitely that the lifetime of the most energetic bubble nuclei is not more than 1 millisecond, and that some of the weaker nuclei that contribute bubbles to minimum-ionizing tracks have lifetimes less than 0.1 millisecond. These results

obtained with the two-inch chamber depend somewhat on the geometry of the chamber and the hydrodynamics of the expansion process, but the general orders of magnitude are probably characteristic of this type of bubble chamber operation. Possibly working liquids with considerably different thermodynamic and mechanical properties will give different results.

The shortness of the lifetime of bubble nuclei happily tends to reduce the background of unwanted ionization events for work with high-intensity pulsed accelerators, but unhappily will probably make counter-controlled expansions for cosmic-ray experiments impossible with this technique. Before undertaking these detailed sensitive-time measurements we tried without success to photograph cosmic-ray tracks by means of counter-controlled expansions.

SUMMARY OF PROPERTIES OF BUBBLE CHAMBERS

As a result of these developments it is now fairly easy to construct a bubble chamber with a sensitive volume of several liters of diethyl ether, iso-pentane, or some similar liquid. For these liquids the density of the sensitive medium is about 0.5 g/cm³ of which 0.1 g/cm³ is hydrogen. Liquid hydrogen chambers of almost 100 cm³ containing 0.05 g/cm³ of hydrogen have also been operated successfully.[2] Chambers with a long dimension of several feet seem feasible with recent techniques. Other liquids of various densities and chemical compositions can be used. A very rough guide for estimating the operating conditions for a given liquid is that the operating pressure is about two-thirds of the critical pressure and the operating temperature lies about two-thirds of the way from the normal boiling temperature to the critical temperature. A more accurate estimate can be obtained with the published formula[1,4] if one takes $n=6$. For diethyl ether and iso-pentane the volume expansion ratio is found to be 1.03. These chambers can be expanded in less than five milliseconds and may have a sensitive time of the order of ten milliseconds. Because the lifetime of the bubble nuclei is so short and the sensitivity increases from zero to a maximum in about 1 millisecond, these chambers can be operated in a region of high radiation background without excessive bubble background. For operation with pulsed beams, the chamber can be expanded just before the beam arrival and pictures can be taken by counter-controlled flashlamps. Since the bubbles grow large enough to photograph in a few microseconds, there are virtually no distortions of the tracks due to motions in the liquid. Another advantage of the chambers is that their operation cycle can be as short as three or four seconds, even for chambers containing a liter of liquid. This makes it possible to accept particles from every pulse of a large accelerator such as the Brookhaven Cosmotron, which delivers a pulse once every five seconds.

For bubble chambers operated with liquids of low atomic number, Coulomb scattering will not interfere seriously with measurements of magnetic curvature in magnetic fields of convenient size. At maximum sensitivity a minimum-ionizing particle in diethyl ether or iso-pentane makes a track containing about 100 bubbles per cm, making possible precision measurements of ranges of particles up to a range of tens of g/cm². Improved photography should make possible the detection of bubbles as small as 0.01 mm in diameter. Then magnetic curvature measurements will be limited principally by optical problems, and it may become possible to measure scattering along a track.

Finally it has been found that the bubble density along a track varies with the ionization of the particle, making the measurement of ionization possible by counting bubbles along a track. Minimum-ionizing particles can produce 100 bubbles per centimeter down to less than 1 bubble per centimeter depending on the temperature of the liquid and the pressure. This fact leads to the interesting possibility that the bubble chamber sensitivity can be adjusted for different experiments, but may give rise to great experimental difficulties in the attempt to calibrate bubble density *versus* ionization density.

We are grateful to the Cosmotron staff at the Brookhaven National Laboratory for making possible our tests of the six-inch bubble chamber, and especially to the cloud chamber group at the Laboratory for their considerable help in carrying out the experiments.

Scattering of 119-Mev Pions by Deuterium*

D. E. NAGLE
Institute for Nuclear Studies and Department of Physics, The University of Chicago, Chicago, Illinois
(Received September 23, 1954)

The angular distributions for the scattering of negative pions by liquid deuterium were measured, using the 119-Mev pion beam of the Chicago cyclotron and scintillation counters. The values of the nonexchange scattering differential cross section in the laboratory system for angles of 45°, 90°, and 135° are 8.50±0.31, 5.31±0.20, and 8.36±0.29 in units 10^{-27} cm² steradian^{-1}. The differential cross sections for the production of photons at the same angles and in the same units are 2.20±0.27, 2.41±0.28, and 3.40±0.38. The total cross section from transmission measurements is $(132\pm7)\times10^{-27}$ cm². These values are somewhat less than the sum of the previously measured cross sections for positive and for negative pions on hydrogen.

INTRODUCTION

THE following reactions are expected between a fast negative pion and a deuteron:

$\pi^-+d\rightarrow\pi^-+d$ (elastic scattering), (1a)

$\pi^-+d\rightarrow\pi^-+n+p$ (inelastic scattering), (1b)

$\pi^-+d\rightarrow\pi^0+n+n$ (charge exchange scattering), (2)

$\pi^-+d\rightarrow n+n$ (absorption), (3)

$\pi^-+d\rightarrow n+n+\gamma$ (radiative absorption). (4)

Five similar reactions may be written by exchanging π^+ for π^- wherever it appears, and exchanging neutrons and protons. The principle of charge symmetry for meson-nucleon interactions implies that each of the above reactions would occur with the same probability as its charge image reaction, except for Coulomb effects.

In the limiting case where the deuteron is regarded as unbound, charge symmetry being assumed, the deuterium cross section is expected to be the sum of the cross sections for positive pions and for negative pions scattered by hydrogen. In fact, the recent transmission experiments indicate that this is approximately true[1,2] but that the deuterium total cross section is less than the sum of the hydrogen cross sections.[3] It is of considerable interest to extend the measurements to the differential cross sections.

Measurements of the differential cross sections for the scattering by hydrogen of 120-Mev pions have been reported.[3] The following describes measurements of the differential cross sections for the scattering of 119-Mev negative pions by deuterium, that is, reactions (1a), (1b), and (2) above. Scintillation counters and a liquid deuterium target were used, in a geometrical arrangement similar to that used for hydrogen.[3] A measurement of the distribution in range of the scattered pions was included, and a measurement of the total cross section.

I. EXPERIMENTAL ARRANGEMENT

The pion beam was produced by 450-Mev protons in the Chicago synchrocyclotron striking a beryllium target at 76-inch radius. Negative pions emitted in the forward direction were bent outward by the fringing magnetic field of the cyclotron, and emerged through a thin aluminum window in the vacuum chamber of the machine. The pions passed through a channel cut through the 12-foot shield surrounding the cyclotron, and on the experimental side of the shield were further deflected by an analyzing magnet. The pions passed through two counters, called the doubles counters, which were 2-inch diameter scintillation counters (No. 1 and No. 2 in Table I) and then entered the deuterium container. A scintillation counter telescope could be placed to detect the scattered particles at any chosen angle. The number of coincidences recorded with the doubles counters measured the number of pions falling on the scatterer. The coincidences of the scattered particle counters with the doubles counters measured the number of particles scattered within the solid angle of the detector telescope. The ratio of these numbers was essentially the fraction of the incident beam which scattered into that solid angle.

II. SCINTILLATION COUNTING EQUIPMENT

The scintillation counters used are listed in Table I. The scintillating liquid was phenylcyclohexane with 3 g per liter of terphenyl and 10 mg per liter of diphenylhexatriene dissolved in it.[4]

The pulses from the photomultiplier anodes were fed through a preamplifier and pulse shaping circuit attached to the counter case. The output pulses were roughly square and about 2.5×10^{-8} sec long. After amplification in two distributed amplifiers they went into coincidence circuits. For the measurement of the scattering of pions (see Fig. 4) pulses from counters No. 1 and No. 2 were brought into a circuit "B," one output of which led to a Hewlett Packard Model 520A tenth microsecond scale of one hundred and another output to a triple coincidence circuit. Pulses from counters No. 5 and No. 4 were brought into coincidence

* Research supported by a joint program of the Office of Naval Research and the U. S. Atomic Energy Commission.
[1] Anderson, Fermi, Martin, Nagle, and Yodh, Phys. Rev. 86, 413 (1952).
[2] Isaacs, Sachs, and Steinberger, Phys. Rev. 85, 803 (1952).
[3] Anderson, Fermi, Martin, and Nagle, Phys. Rev. 91, 155 (1953).
[4] H. Kallmann and M. Furst, Phys. Rev. 81, 853 (1951).

in a second circuit "D," and the output into the triples circuit. The output pulses of "B," "C," and "D" were each 5×10^{-8} sec long. The triples circuit in turn drove another Hewlett Packard tenth microsecond scaler. The Hewlett-Packard scalers drove Atomic Instrument scalers. The overall resolving time of the system was about 2.5×10^{-8} sec and the dead time, principally due to the scalers, was 0.1 microsecond. The doubles counting rate was kept nearly constant throughout these experiments at 250 000/min. Because of the duty factor of the cyclotron (about 60), this amounted to a loss in doubles counts in the scaler of about 2.5 percent. The equipment was designed by Anderson, Glicksman, and Martin, and has been described in detail by them.[5]

III. DEUTERIUM CONTAINER

The deuterium was condensed from the gas using liquid hydrogen as the refrigerant. A cross-sectional view of the apparatus is shown in Fig. 1; it is essentially a continuously pumped metal dewar, holding about six liters of liquid hydrogen. Evaporating hydrogen gas cools one set of coils of a heat exchanger; a second set of coils in contact with the first cools the incoming deuterium gas. Finally a third coil wound around the hydrogen reservoir condenses the deuterium. The liquid collects in the deuterium reservoir and in the scattering chamber. The scattering chamber is a brass cylindrical container 4.250 inches inner diameter and 5 inches high at room temperature. The piping is arranged so that by closing the valve separating the heat exchanger from the deuterium reservoir, the liquid deuterium will stay in either the reservoir or the scattering chamber, whichever is colder; since either can be heated electrically, the liquid can be switched from one to the other chamber. In this way observations can be made of the fraction of pions scattered with and without liquid in the scattering chamber. Capacitance type level indicators are used to display the level of the liquid deuterium. The mean length in the deuterium is 4.10 inches, including a profile factor (Sec. IV).

TABLE I. Characteristics of the scintillating counters.

Counter No.	Scintillator Dimensions	Material	Photomultiplier type
1	2 inches diam × $\frac{1}{8}$ inch	plastic	RCA 5819
2	2 inches diam × $\frac{1}{16}$ inch	liquid	RCA 5819
3	4 inches × 6 inches × $\frac{1}{4}$ inch	liquid	RCA C7157
4	4 inches × 6 inches × $\frac{1}{4}$ inch	liquid	RCA C7157
5	4 inches × 6 inches × $\frac{1}{4}$ inch	liquid	RCA C7157
6 and 7	8 inches diam × $\frac{3}{4}$ inch	liquid	RCA C7157

[5] Anderson, Martin, and Glicksman, Proc. Natl. Electronics Conf. 9, 483 (1953).

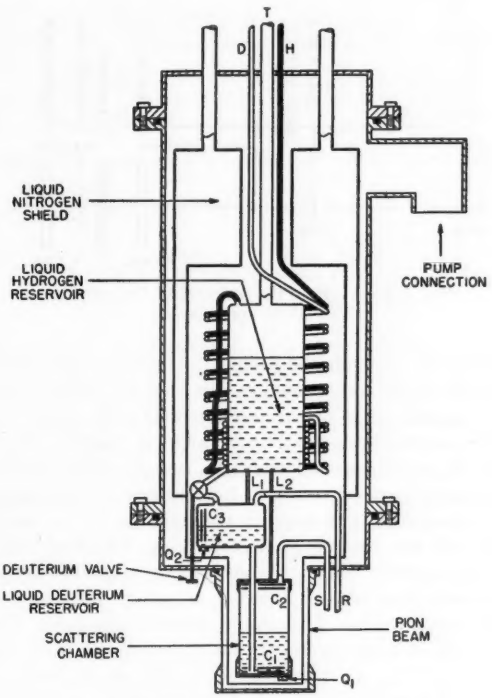


FIG. 1. Liquid deuterium container—section through a vertical axis.

The temperature of the liquid deuterium, estimated from the vapor pressure, is 24°K. If one uses the data of Woolley *et al.*,[6] the net number of atoms/cm² (liquid less gas) is 4.813×10^{23}/cm² (1 ± 0.018). Of these 98.9 percent are deuterium and 1.1 percent are hydrogen atoms, so that the number of d atoms/cm² is 4.76×10^{23}.[7]

The pion beam on entering or leaving the target must pass through metal walls of thickness 0.22 g/cm² of Al and 0.12 g/cm² of brass.

IV. ENERGY, COMPOSITION, AND PROFILE OF THE PRIMARY BEAM

The energy of the beam has been studied at various times by measurements of the magnetic rigidity, of the Panofsky effect, and of the range curve.[5] The results agree well; only the range curve is here presented.

For this purpose, the deuterium target was removed from the beam and the counters placed in line as shown in Fig. 2. The double coincidence of counters No. 1 and No. 2, and the quintuple coincidences were recorded. Figure 3 is a plot of the first difference of Q/D against thickness of copper absorber. The curve is a smooth

[6] Woolley, Scott, and Brickwedde, J. Research Natl. Bur. Standards 41, 379 (1948).

[7] Professor M. G. Inghram very kindly made this measurement for us.

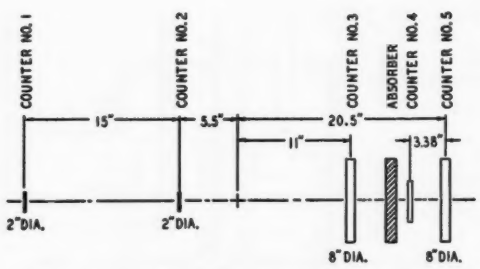


FIG. 2. Plan view of the geometry of the counters for taking range curves.

one drawn through the points; it shows a large peak due to pions and a smaller one due to muons. The uncorrected range of the pions is 47.2 g of copper; to which is added the copper equivalent of counters No. 3 and No. 4, plus a wall of No. 5, namely 5.1 g. To get the range from the center of the deuterium, there must be subtracted the copper equivalent of the walls and half the deuterium, namely 1.6 g. The net range, 50.7 g, corresponds to a kinetic energy of 119 Mev if one uses the tables of Aron[8] of ranges for protons and adopts the ratio 6.72 of proton mass to pion mass. The fraction of pions in the beam is estimated from the curve as 0.93. To measure the profile of the beam, a piece of plastic scintillator of cross section $\frac{3}{16}$ inch by $\frac{5}{16}$ inch was mounted on a Lucite light pipe leading to a 5819 photomultiplier. The coincidence of this counter with the doubles of the counters No. 1 and No. 2 described above were recorded, and the ratio of triple coincidences to doubles of one and two plotted as a function of position of the third. This measurement was made at the beginning of each cyclotron run. It showed a beam whose intensity for a traverse across the horizontal plane showed a plateau shaped curve which dipped sharply at about $\frac{3}{4}$-inch radius from the target location. Because of this extent of the beam the number of atoms of deuterium per cm² quoted in Sec. III includes a factor of 0.97, the so-called profile factor.

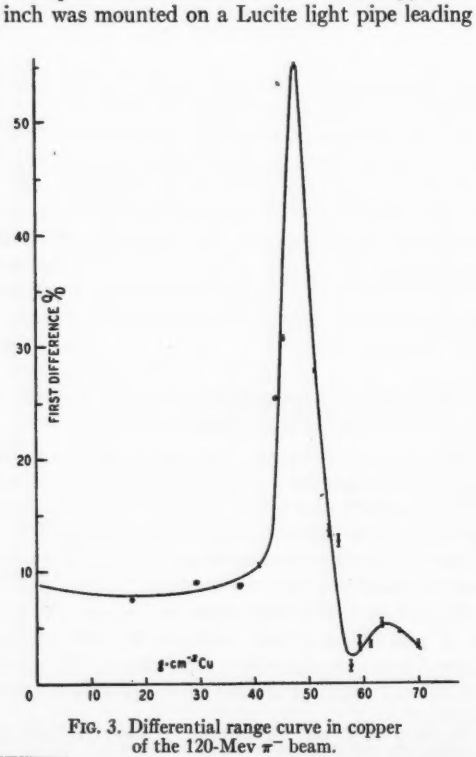


FIG. 3. Differential range curve in copper of the 120-Mev π^- beam.

[8] W. A. Aron, University of California Radiation Laboratory Report UCRL-1325, 1951 (unpublished).

V. MEASUREMENT OF THE CROSS SECTION FOR PRODUCTION OF PHOTONS

The charge-exchange scattering leads to production of photons from the decay of the neutral meson. There are in addition a much smaller number of photons from the radiative absorption. Figure 4 shows a plan view of the arrangement used to count photons. An

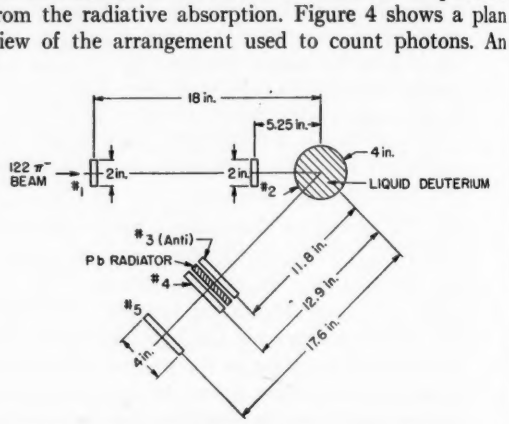


FIG. 4. Geometry of the counters for measurements of scattering.

event typical of those recorded was that a pion in the beam passed through counters No. 1 and No. 2, scattered with charge exchange in the deuterium, producing a neutral meson which decayed to produce a photon traveling in the line of counters Nos. 3, 4, and 5. Such a photon, if it materialized in the lead radiator (7.33 g cm^{-2}), produced an electron or pair which gave a coincidence pulse between 4, 5, 1, and 2, but no pulse in No. 3, which was connected in anticoincidence. The ratio Q/D of quadruple coincidences to doubles of No. 1 and No. 2 is a measure of the number of photons produced within the solid angle of the detector telescope. Table II lists the observed values of Q/D for three angular positions of the photon counter. The data were taken on two occasions, labelled Run A and Run B in the table. In Run B additional lead shielding was placed near the steering magnet, which helped to reduce the background. The errors listed are statistical only. The difference of the Q/D values with and without liquid deuterium, the net value, is listed in the last column. The net values for both runs are combined

according to their statistical weights, and listed at the bottom of the table.

The sensitivity of the detecting counters is calculated in a manner similar to that used in previous experiments on hydrogen.[3] The introduction of the anticoincidence technique in the present experiment, and slight differences in the geometry necessitate changes in the details of the computation. However, the resulting cross sections for photon production may be compared with the hydrogen cross sections without much disturbance from any inadequacy in the calculation.

Some of the photons materialize before reaching the anticoincidence counter and consequently are not counted. This materialization (cross section for pair production × atoms/cm²) is the sum of contributions from the liquid deuterium, dewar walls, the anticoincidence counter, and the internal conversion of the photon, giving a net value 0.031 for 90-Mev photons, or 0.032 for 135-Mev photons. The materialization after the anticoincidence counter is due to the 7.3 g of lead plus a small contribution from counter No. 4. It has the value 0.677 for 90-Mev photons and 0.731 for 135-Mev photons.

TABLE II. Observed values of $10^6 G/D$ for the photon counter.

Angle	With deuterium	Without deuterium	Net
	Run *A*		
45°	94.1±4.0	52.7±5.0	41.4±6.4
90°	150.7±4.3	103.1±3.8	47.6±5.7
135°	128.1±4.6	67.3±3.7	60.8±5.9
	Run *B*		
45°	68.3±2.9	26.8±1.4	41.5±3.2
90°	61.9±2.7	20.0±1.3	41.9±3.0
135°	92.2±3.0	31.5±2.0	60.7±3.6
	Runs *A* and *B* combined		
45°			41.5±2.9
90°			43.1±2.7
135°			60.7±3.1

The efficiency ϵ_γ of the counter for detecting a photon is computed, taking into account: (a) materialization before and after anticoincidence counter, (b) absorption of the electrons in the lead, and (c) the geometrical effect due to multiple scattering of the electrons in the lead. This gives a very slightly different effect at 90° than at 45° or 135°, because of the difference in the effective lengths of the source of photons. The results are listed in Table III. The efficiencies are lower than in the previous arrangement[3] because (a) the prematerialization now lowers instead of raising the efficiency, and (b) the geometrical effects of multiple scattering are smaller. The anticoincidence arrangement has the advantage of a rather low background. The net Q/D for photons is related to the cross section for photon production by the equation:

$$\left(\frac{Q}{D}\right)_\gamma = \frac{d\sigma_D}{d\omega}\left[1+0.011\frac{d\sigma_H}{d\omega}\Big/\frac{d\sigma_D}{d\omega}\right]C\epsilon_\gamma.$$

TABLE III. Efficiency for photons.

Angle of scattering	90 Mev	135 Mev
45° or 135°	0.470	0.556
90°	0.466	0.556

The second factor on the right is the correction due to hydrogen atoms and has the value 1.012. C contains the following factors: the solid angle, the number of D atoms/cm², the fraction of pions in the beam, the mean transmission excluding scattering to the center of the deuterium, and mean number of double coincidences per doubles scaler count. In that order the factors are $C=0.0772\times4.76\times10^{22}/\text{cm}^2\times0.93\times0.992\times1.025$. The values of $(d\sigma_D/d\omega)_\gamma$ and the previous values[3] of $(d\sigma_H/d\omega)_\gamma$ are listed in Table IV. The errors in $(d\sigma_D/d\omega)_\gamma$ are compounded of the statistical error plus 10 percent of the cross section, which is an estimate of the uncertainties arising from the efficiency calculation, pion fraction, doubles loss, and number of D atoms/cm².

VI. SCATTERING WITHOUT EXCHANGE OF CHARGE

The cross sections for scattering without charge exchange were also measured with the arrangement of Fig. 4, except that counter No. 3 was disconnected from the circuits. The observed ratios of quadruples to doubles at angles of 45, 90, and 135 degrees appears in Table V. Here three runs are combined.

The Q/D values are related to the cross sections by the equation

$$\frac{Q}{D}=\frac{d\sigma_D}{d\omega}\left[1+\frac{n_H\sigma_H}{n_D\sigma_D}\right]C\epsilon_\pi+\Gamma.$$

The second factor is here 1.001. C has the value given in Sec. V. The efficiency for counting pions, ϵ_π, differs from unity because of roughly equal losses from multiple scattering and nuclear absorption. The values of ϵ_π are 0.922, 0.919, and 0.915 at 45°, 90°, and 135° respectively. Γ is a small contribution to Q/D from photons converted in counters Nos. 3 and 4, which is 0.031 times the photon flux δ (Sec. V). The computed values of $d\sigma_D/d\omega$ are listed in Table VI, and, for com-

TABLE IV. Differential cross sections for the production of photons (laboratory system). Units are 10^{-27} cm²/sterad.

Angle	$(d\sigma_D/d\omega)_\gamma$	$(d\sigma_H/d\omega)_\gamma$
45°	2.20±0.27	2.64±0.36
90°	2.41±0.28	3.08±0.37
135°	3.40±0.38	4.53±0.51

TABLE V. Observed values of $10^6 Q/D$ for scattering of negative pions without charge exchange.

Angle	With deuterium	Without deuterium	Net (Without Pb)
	Run *A*		
45°	553.3±10.0	249.0±8.4	304.3±13.0
90°	307.8± 6.0	119.7±3.8	188.1± 7.1
135°	424.6± 8.8	149.1±5.6	275.5±10.4
	Run *B*		
45°	438.3± 7.4	165.4±5.3	272.9± 9.1
90°	218.3± 5.2	56.1±3.1	162.2± 6.1
135°	360.9± 6.6	91.8±3.9	269.1± 7.7
	Run *C*		
45°	476.6± 4.9	195.8±4.9	280.8± 6.8
90°	238.6± 6.5	60.6±3.2	178.0± 7.2
135°	378.8± 6.1	97.6±4.4	281.2± 7.5
	Runs *A*, *B*, and *C* combined		
45°			282 ± 5.0
90°			175 ± 3.9
135°			275 ± 4.8

parison, the sum of the hydrogen ordinary scattering cross sections for positive and for negative pions. The errors are the statistical one compounded with 3 percent, which is our estimate of the errors in C and ϵ_π.

VII. ENERGIES OF THE SCATTERED PIONS

An attempt was made to find the energy distributions of the scattered pions by taking range curves. Figure 5 shows the experimental arrangement. Double coincidences of counters No. 1 and No. 2, quintuple coincidences were recorded. Counter No. 3 serves to reduce the background due to neutrons or photons, and counter No. 4 defines the solid angle of the counter telescope. Counter No. 5 is chosen to have a large solid angle in order to effectively eliminate the loss of coincidences due to multiple scattering in the absorbers.

The observed values of Q/D are presented in Table VII in various absorber thicknesses at the three laboratory angles of 45°, 90°, and 135°. The values given are net (deuterium in minus deuterium out) and include a small correction made for the nuclear absorption in copper.[9,10] In Fig. 6 is plotted the first difference of Q/D divided by the difference in absorber thickness, *vs* the mean absorber thickness. The smooth curves of Fig. 6 have been computed as follows: from the 0° range curve, Fig. 3, the mean energy of pions at the center of the deuterium was estimated as 119 Mev. The pion peak at 0° is approximately gaussian in shape with a root mean square width about 2.5 g of copper. The energy of elastically scattered pions can be calculated at each scattering angle from simple kinematics, as can the energy of a pion scattered from a free proton. An approximate fit is obtained by assuming the points are represented by the sum of two Gaussians centered about the elastic energy and the free proton energy. The width of the elastic Gaussian is taken as 2.5 g in the 0° case, since the effects of straggling, the deuterium, and the walls do not appreciably broaden the peak. The width of the free-proton Gaussian and the amplitudes of the Gaussians are treated as parameters in a least-squares analysis. The results are the smooth curves plotted. The widths of the free-proton Gaussian are 4 g, 4 g, and 5 g at the three angles. The percent elastic scattering at the angles of 45°, 90°, and 135° are 30 percent, 18 percent, and 21 percent by this method, giving 2.6 mb/steradian, 0.95 mb/steradian, and 1.8 mb/steradian for the differential cross sections for elastic scattering at the corresponding three angles. The barycentric cross sections and angles are 2.2, 0.95, and 2.2 mb/steradian at 50.3, 97.3, and 140.0° respectively.

The method is very rough, especially at forward angles since the elastic and the "free-proton" peaks become too close to allow resolving the peaks. The

TABLE VI. Differential cross sections in the laboratory system for the scattering of negative pions without charge exchange. Units are 10^{-27} cm²/sterad.

Laboratory angle	$d\sigma_D/d\omega$	$d\sigma_{H^+}/d\omega + d\sigma_{H^-}/d\omega$
45°	8.50±0.31	7.22±1.58
90°	5.31±0.20	5.94±1.11
135°	8.36±0.29	11.57±1.24

TABLE VII. Observed values of Q/D for the scattered pions.

Angle	Absorber thickness grams of copper	$10^6 Q/D$ (net)
45°	0	249.1±11.0
	11.5	219.7± 8.8
	20	167.8±11.0
	24.1	169.7±10.8
	28.3	144.9± 8.1
	32.5	119.5± 8.7
	36.7	82.3± 8.4
	41.0	27.6± 6.6
	45.2	5.7± 4.7
90°	0	169.4± 9.4
	8.47	137.6± 7.8
	15.8	104.5±11.7
	20.0	76.5± 5.5
	24.2	54.9± 4.6
	28.4	22.9± 4.6
	32.6	11.2± 2.9
	36.8	3.8± 2.6
	41.0	3.8± 1.8
135°	0	272 ±13
	4.2	250.4±13.5
	8.5	187.0±11.0
	11.5	134.0± 8.5
	15.8	77.0± 6.5
	20.0	52.4± 5.1
	24.2	19.2± 4.0
	28.4	6.8± 3.4
	32.6	5.5± 2.5

[9] Chedester, Isaacs, Sachs, and Steinberger, Phys. Rev. 82, 958 (1951).
[10] Martin, Anderson, and Yodh, Phys. Rev. 85, 486 (1952).

points of Fig. 6 indicate that the elastic scattering cross section is small at 90° and 135° and suggest that it is somewhat larger at 45°. There appears to be a considerable number (about 25 percent of the total) of pions at all angles with energies lower than the free nucleon energy.

VIII. TOTAL CROSS SECTIONS

The total cross section has been obtained by transmission measurements. The old value from heavy water-light water difference was 128±11 mb for 120 Mev negative pions.[1] The value from transmission in liquid deuterium gave 132±7 mb. These have been corrected for reaction products recorded by the final counters.

Integration of the differential cross sections gives for the cross section for scattering without exchange of charge 96±5 mb and for charge exchange 14±2 mb. The radiative absorption may be estimated as 1 mb. The absorption process $(\pi^+ + d \rightarrow p + p)$ has been studied by Stadler[11] in this laboratory. He finds the total cross section to be 7.4±1.2 mb for 91-Mev pions and 9.7±1.2 mb for 114-Mev pions. Invoking charge symmetry and extrapolating his values to 119 Mev we adopt 11±2 mb for the absorption cross section. The radiative absorption cross section may be estimated as 1 mb. The sum of cross section for the four processes gives 122±6 mb, in good agreement with the transmission values. For comparison, the sum of the total cross sections in hydrogen for positive and negative pions gives 130±11 mb from the integration of the angular distributions and 147±14 mb from the transmission experiments.[1] The integrated values are believed more reliable.

IX. DISCUSSION

The experimental results of Tables IV and VI indicate that simple additivity of the cross sections for free nucleons is a fair approximation to the deuteron cross section at this energy. The deuteron cross sections however are less than the sum of the hydrogen values, the deviations ranging from ten to thirty percent. (One exception is the 45° ordinary scattering.)

Several theoretical discussions have been given of the relationship of the pion-deuteron and the pion-hydrogen scattering processes,[12,13] making use of the impulse approximation.[14] Fernbach *et al.*[12] have made extensive use of the closure approximation. In this approximation particularly simple expressions are obtained, which can be evaluated using the pion-hydrogen phase shifts. The cross section for charge-exchange

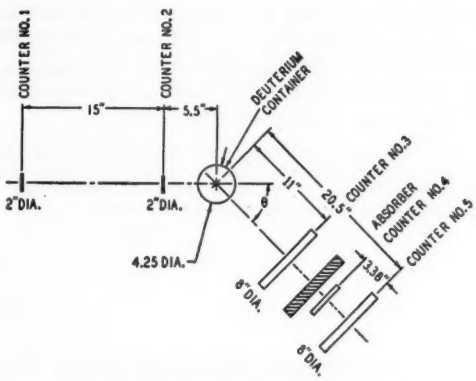


FIG. 5. Experimental arrangement for taking ranges of the scattered pions.

scattering in this approximation is the hydrogen charge-exchange cross section multiplied by a factor depending on the phase shifts and the scattering angle. The angular distributions for photons calculated from this expression agree, within the experimental error, with the observed cross sections.

However the ordinary-scattering cross sections in the closure approximation at this energy should be greater than the sum for hydrogen, contrary to observation.

Brueckner[13] has shown that, in the case of elastic scattering, a calculation based on a model, which takes into account multiple scattering of the pion from the

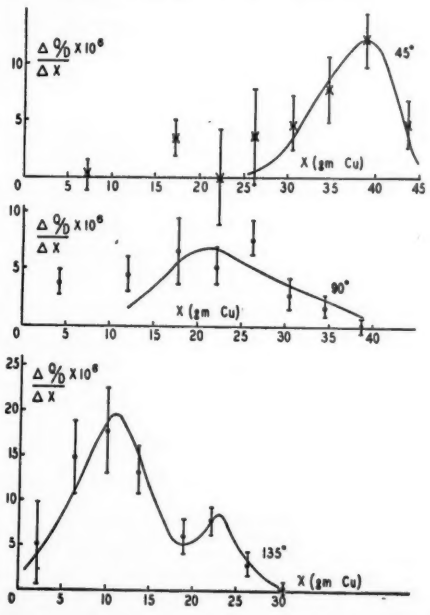


FIG. 6. Differential range curves for the scattered pions. The effective ranges corresponding to elastic scattering at 45°, 90°, and 135° are 40, 32, and 23 g/cm², respectively. For free-nucleon scattering the corresponding effective ranges are 36, 21, and 11 g/cm².

[11] H. Stadler, Ph.D. Thesis, University of Chicago, 1954 [Phys. Rev. **96**, 496 (1954)].

[12] Fernbach, Green, and Watson, Phys. Rev. **84**, 1084 (1951); Fernbach, Green, and Watson, Phys. Rev. **89**, 834 (1953); Thomas A. Green, Phys. Rev. **90**, 161 (1953).

[13] K. A. Brueckner, Phys. Rev. **90**, 715 (1953); K. A. Brueckner, Phys. Rev. **89**, 834 (1953).

[14] G. F. Chew, Phys. Rev. **80**, 196 (1950); G. F. Chew and M. L. Goldberger, Phys. Rev. **87**, 778 (1952).

two nucleons, can lead to a cross section smaller than given by the previous calculations based on the impulse approximation, especially where the phase shift α_{33} is large, as is here the case. The experiments suggest that such effects may be important.[15]

[15] Noteworthy in this connection are the experiments of Arase, Goldhaber, and Goldhaber, Phys. Rev. **90**, 160 (1953). Their cross sections for elastic scattering are much smaller than the earlier calculation (see reference 12) would predict.

X. ACKNOWLEDGMENTS

I wish to thank Professors Earl A. Long and Lothar Meyer for much helpful advice and assistance in the construction of the deuterium target, Professor Herbert L. Anderson for the use of his counting equipment, and Charles Cohn, Henry Stadler, and Willard Skolnik for help in taking data. I wish also to acknowledge the help of a grant from the Guggenheim Foundation.

PHYSICAL REVIEW VOLUME 97, NUMBER 2 JANUARY 15, 1955

Landau Distribution and Density Effect at High Gas Pressures*

E. D. PALMATIER, J. T. MEERS,† AND C. M. ASKEY‡
Department of Physics, University of North Carolina, Chapel Hill, North Carolina
(Received September 27, 1954)

Studies of the ionization energy losses of fast μ mesons have been carried out in a specially constructed proportional counter with argon gas at pressures up to 40 atmospheres. With increasing pressure, a strong reduction in the relativistic rise was observed; the extent of the reduction being at least as great as that expected from Sternheimer's calculation of the density effect. Furthermore, the distribution in energy losses, which at low pressures is much wider than expected from theoretical considerations, was observed to become narrower with increasing pressure. At our highest pressure, however, the energy loss distribution was still considerably wider than predicted by the Landau theory.

I. INTRODUCTION

PUBLISHED studies of the ionization energy losses of fast μ mesons confirm the existence of the relativistic rise in dispersed media[1–4] as well as the modifications of this rise (the density effect) caused by the dielectric polarization, when the medium is condensed.[5–10] Some observations also indicate the onset of the density effect, even at low pressures in gases, provided the meson energy is sufficiently high.[2,4,11]

The general features of the distribution in energy loss, as developed by Landau[12] have also been confirmed, for both dispersed and condensed media. In general, however, with small thicknesses of absorber, the widths of the observed distributions are found to be much greater than those predicted by the theory of Landau or by the modified theory of Blunck and Leisegang.[13] It should be noted, for the case of low-energy electrons traversing thin foils that a similar disagreement was reported although recent work appears to have eliminated the discrepancy.[14,15] The half-maximum widths observed with large thicknesses of absorber are generally in much better agreement with theory.[5,9]

This report deals with a study of μ-meson ionization energy losses, carried out in argon gas at pressures up to 40 atmospheres, in an attempt to observe the onset and the development of the density effect and the behavior of the Landau distribution in the transition region between the dispersed and condensed phases of matter.

II. DESCRIPTION OF APPARATUS

The apparatus used in this work, in particular the electronic and recording portions, was quite similar to that used in our previous studies,[3] in this laboratory. It consisted essentially of four parallel Geiger counter telescopes, with enough lead absorber to insure that four well collimated beams of mesons were obtained; two proportional counters located within the telescopes; lead absorber and range trays; plus the required re-

* This research was entirely supported by the U. S. Atomic Energy Commission. This report is based in part on two Ph.D. dissertations, submitted by two of the authors (C.M.A. and J.T.M.) to the Graduate School of the University of North Carolina and in part on separate work conducted on the U.S.A.E.C. program.

† Now at National Carbon Research Laboratory, Cleveland, Ohio.

‡ Now at the College of Charleston, Charleston, South Carolina.

[1] Becker, Chanson, Nageotte, Treille, Price, and Rothwell, Proc. Phys. Soc. (London) **A65**, 437 (1952); **A66**, 167 (1953).

[2] Parry, Rathgeber, and Rouse, Proc. Phys. Soc. (London) **A66**, 541 (1953).

[3] J. E. Kupperian and E. D. Palmatier, Phys. Rev. **91**, 1186 (1953).

[4] Ghosh, Jones, and Wilson, Proc. Phys. Soc. (London) **A65**, 68 (1952).

[5] W. L. Whittemore and J. C. Street, Phys. Rev. **76**, 1786 (1949).

[6] E. Pickup and L. Voyvodic, Phys. Rev. **80**, 89 (1950).

[7] T. Bowen and F. X. Roser, Phys. Rev. **83**, 689 (1951).

[8] M. M. Shapiro and B. Stiller, Phys. Rev. **87**, 682 (1952).

[9] A. Hudson and R. Hofstadter, Phys. Rev. **88**, 589 (1952).

[10] B. Stiller and M. M. Shapiro, Phys. Rev. **92**, 735 (1953).

[11] Ghosh, Jones, and Wilson, Proc. Phys. Soc. (London) **A67**, 331 (1954).

[12] L. Landau, J. Phys. (U.S.S.R.) **8**, 201 (1944); K. R. Symon in *High Energy Particles* by B. Rossi (Prentice-Hall Publications, New York, 1952).

[13] O. Blunck and S. Leisegang, Z. Physik **128**, 500 (1950).

[14] F. Kahl and R. D. Birkhoff, Phys. Rev. **91**, 505 (1953).

[15] E. T. Hungerford and R. D. Birkhoff, Phys. Rev. **95**, 6 (1954).

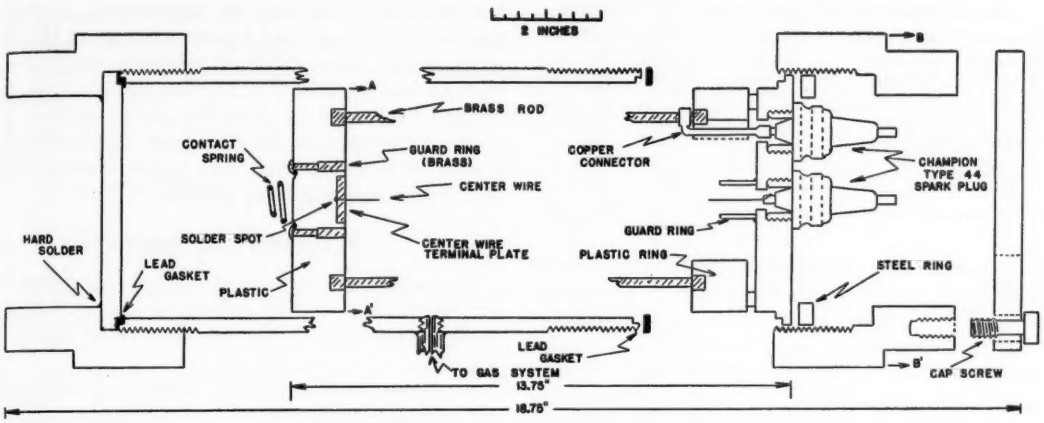


FIG. 1. Schematic drawing of high-pressure proportional counter.

cording circuits. The following main modifications were made to our original apparatus.

(a) The dimensions of each of the four telescopes were set at 2 in.×1.25 in.×13 in., the greatest dimension lying in the vertical. This gave a total counting rate of 50 mesons per hour without destroying the collimation of the mesons or introducing a serious spread in the meson path length within the proportional counter. Thus for this arrangement, the extreme path lengths through the counter differed from the mean path length by less than three percent.

(b) The amount of lead absorber was greatly increased to attain greater momentum limits. The dimensions of the absorbers are given in Table I.

(c) An extensive set of guard counters was used to eliminate the effect of showers falling upon the apparatus.

(d) The lower proportional counter, which was run at a pressure of 2 atmospheres throughout the whole of the experiment, was one of the counters used in our previous experiment, modified by the substitution of a guard ring seal (Stupakoff Type 96.0019) in place of the previous seal. Although the type of seal used possesses excellent electrical insulation properties, it will not withstand a pressure differential of much more than two and one-half atmospheres.

(e) The upper proportional counter in which the high pressures were used, was constructed as shown in Figs. 1 and 2. It consisted of a core, which carried all the electrical components of the counter, and of a thick-walled steel shell into which the core fitted. The core was composed of six brass rods (forming a cylindrical equipotential surface concentric with the center wire which served as the outer electrical conductor), a center wire of tungsten and two end supports of Lucite, all of which were carried by a heavy steel end plate. The electrical connections to the outer conductor and the center wire, which were run at negative and ground potential, respectively, were carried through the steel end plate by means of two Champion Type 44 Diesel spark plugs. Guard rings were inserted as shown, one making direct contact with the end plate, the other making contact with the shell by means of the contact spring. The shell was grounded during the course of the experiment. A lead washer was used with the end plate. No gas leaks were observed with the counter at 120 atmospheres pressure during a 12-hour underwater test and excellent proportional counter pulses

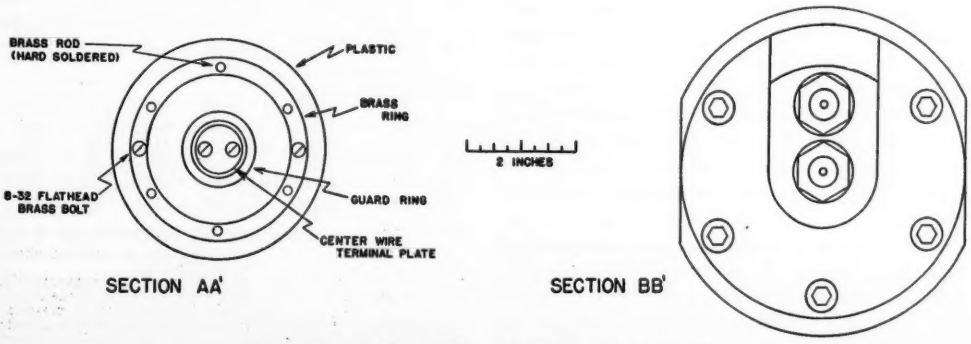


FIG. 2. Cross sections of high-pressure proportional counter.

TABLE I. Dimensions and energy limits for lead absorbers.

Designation of meson group	Absorber thickness (g cm^{-2} Pb)	Energy band (Bev)	Average energy (Bev)	Effective momentum (Bev/c)
A	120– 470	0.165–0.55	0.36	0.45
B	470– 930	0.55 –1.13	0.84	0.93
C	930–2590	1.13 –3.45	2.3	2.4
D	2590 or greater	3.45	13.0	13.0

were observed at a potential of 14 000 volts with an argon filling at 55 atmospheres pressure. No tests of this counter were run at higher pressures or voltages.

An important feature of the counter was the center wire which was constructed of 0.0007-in. diameter tungsten wire. Its positioning was achieved by threading through guide holes as shown in the figure and securing it in position by small wedges located well within the field-free region at the ends of the chamber. When the wire had been located within the core, the latter was placed in a bell jar and the wire was flashed in a vacuum for two seconds at 80 volts dc. This flashing procedure which raised the temperature of the wire to around 2000°C, according to our resistance measurements, was found to be necessary and quite effective in producing a center wire which operated consistently. Immediately upon flashing, the core was inserted into the shell and the counter was sealed.

III. THE OPERATING CONDITIONS

During the preliminary runs it was found that our previous practice of using tank CO_2 along with the argon in the high pressure counter caused electron attachment at the higher pressures (40 atmospheres) used in this study. This effect was quite noticeable by the wide spread introduced into the pulse height distribution curve and by the dependence of the pulse height distribution upon the location of the proportional counter with reference to the telescope, (upon the location of the meson path through the proportional counter). Whether this effect arose from the behavior of CO_2 at higher pressures or whether it was due to some impurity contained in our CO_2, we cannot say; it was found to be impossible to prevent this trouble by passage of the CO_2 over hot calcium turnings. It was found possible, however, to eliminate the effect to the extent that no pulse height dependence upon the location of the meson path could be observed (other than that expected from purely geometrical considerations) by completely eliminating the CO_2 and using only Linde tank argon purified by passage over hot calcium. This latter filling was used throughout with both counters.

During the course of the experiment, the standard tests and checks were performed to ensure the quality of the Geiger counter efficiencies and circuit linearities as well as the constancy of the operating voltages and to reduce spurious events to a minimum. In Table II the operating conditions are summarized. It will be observed that Runs 4 and 5 were performed at the same pressure but under different operating conditions of voltages and amplification. The fact that the operating voltage was varied by about 7.5 percent without any observable change in the histogram characteristics is considered to indicate that truly proportional operation of the apparatus was being achieved.

IV. TREATMENT OF THE DATA

The preliminary step in the analysis of the data was the rejection of certain spurious events which were readily recognizable and fell into three classes as follows:

(a) Shower events. These events were identifiable not only by the triggering of the guard tray but also by the simultaneous large outputs from both counters.[3] We believe that no appreciable error arises from a failure to detect such events.

(b) Events in nonlinear regions of the apparatus. Since the apparatus was found to be linear up to outputs of around 4 times the most probable pulse height output in all cases, and no events were used which were greater than 3.0 times the most probable, no errors will be encountered with this type of event.

(c) Spurious events of very small magnitude. A small number of small events were encountered which arose from a spurious triggering of our telescope combined with a small inefficiency in the clamping tube of the pulse lengthener circuit. These events were readily discernible since they only occurred with the *A* group of mesons (higher-order telescope coincidences were required with the other groups of mesons), since they always appeared as pulses of constant and small size simultaneously in both proportional counters, and since they always lay well below the low-energy end of the Landau pulse height distribution. These criteria all permit positive identification of such events. All other events were considered to be true events and were included in the histograms of the data.

TABLE II. Summary of experimental conditions.

Run	Pressure[a] (atmos)	Voltage	Amplifier gain	No. of mesons in group *A*	*B*	*C*	*D*
Low-pressure counter							
1	2	2130	4×10^4	565	552	1055	637
High-pressure counter							
2	40	9700	1×10^4	147	154	340	182
3	40	9700	1×10^4	148	151	303	197
4	21	7000	0.9×10^4	155	182	288	210
5	21	6500	3.4×10^4	99	99	210	218
6	7	3950	2.0×10^4	133	127	252	146
7	1.25	2100	4×10^4	116[b]	82[b]	219	146

[a] Pressure corrected to 0°C and slight correction applied where required to account for behavior of a Van der Waals gas.

[b] In this run a slight failure occurred in the first range tray. This caused about 20 percent of the *B* range mesons to appear in the *A* group. As a consequence, this group, while useful for estimating the pressure behavior of the widths of the distributions, is of no use whatsoever in a good measurement of the relativistic rise and has not been included in the results of Fig. 7.

In the determination of the main characteristics of the histograms, namely the most probable pulse height and the width at half-maximum, where the latter is defined as the ratio

$$\frac{\text{full width of distribution at half-height}}{\text{most probable value of specific ionization}},$$

the two following methods of analysis were used.

(a) Beherens method. This is an adaptation of the Gauss method of maximum likelihood[16] and is described fully elsewhere.[1]

(b) Quantile (or percentile) method. The details of this method are discussed in standard statistical treatises[17]; we shall discuss here only the application of the method to histograms of the type considered in in this study.

If x represents pulse height and $f(x)$ represents the normalized population from which the sample distribution is to be obtained, the quantile $q(x_0)$ of the distribution associated with a given pulse height x_0 is the fraction of events satisfying the requirement $x \leq x_0$. Thus

$$q(x_0)\int_v^u f(x)dx = \int_v^{x_0} f(x)dx,$$

where u and v are the upper and lower limits, respectively, of the range of x. While the value of v is a property of the Landau distribution, that of u must be chosen arbitrarily. We have truncated our histograms at u equal to 2.4 times the most probable pulse height. This procedure is quite convenient and has been used by previous workers.[1]

When a value of $q(b)$ has been chosen, for example, for the most probable pulse height (b), and the quantile procedure is applied to a sample distribution containing n events to determine b for the sample, the standard deviation of the latter quantity is given by

$$[q(b)p(b)/nf^2]^{\frac{1}{2}},$$

where $p(b)=1-q(b)$ and f (the ordinate of the parent distribution at x_0) is to be taken as the frequency per unit interval at x_0 in the sample distribution,[17] normalized over the range v to u.

With a more complex quantity such as the half-width of the distribution, the procedure is as follows. If b, as before, represents the most probable pulse height and the half-maximum values of x are denoted by a and c for the lower and upper values, respectively, then the half-width $(c-a)/b$ possesses a standard deviation which is a function of the variance and covariance of these related quantities and is given by the square root of the following expression:

$$\frac{\text{var}c}{b^2}+\frac{\text{var}a}{b^2}+\frac{(c-a)^2\,\text{var}b}{b^4}-\frac{2\,\text{cov}(ac)}{b^2}+\frac{2(c-a)}{b^3}\{\text{cov}(ab)-\text{cov}(bc)\},$$

where

$$\text{var}x=\frac{q(x)p(x)}{nf^2(x)},$$

and

$$\text{cov}(x_1,x_2)=\frac{q(x_1)p(x_2)}{nf(x_1)f(x_2)},$$

where

$$x_1<x_2.$$

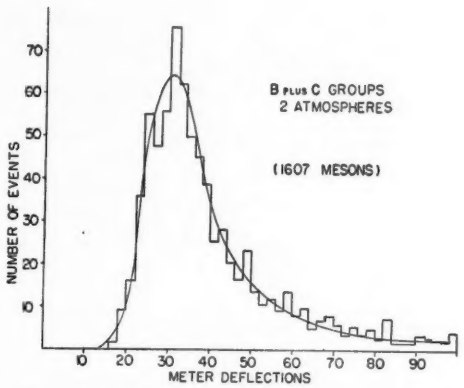


FIG. 3. Basic curve used to fit data and histogram of combined groups from which it was derived.

If the three quantities a, b, and c had been independent of one another, the terms containing the expressions for covariance would have disappeared and the expression for the standard deviation would reduce to the usual form.[18] The same remarks as given previously apply here with respect to the choice of f and it is well to note that care must be taken to obtain the correct selection of p and q in the expressions for the covariance.[17]

Before either of these methods can be applied, a function must be chosen with which to obtain a best fit of the data. The curve used with the Beherens method was obtained from the combined B and C groups for Run 1 (1607 mesons) and was determined by use of the probit method[16] for the lower half of the histogram and a trial and error adjustment to a least-square fit for the upper half. The histogram and our basic curve are shown in Fig. 3. From a study of this curve as well as of all the histograms available to us from the published literature and our own results, it has been possible to determine that the best values for the

[16] N. Arley and K. R. Buch, *Introduction to the Theory of Probability and Statistics* (John Wiley and Sons, Inc., New York, 1950), p. 142.

[17] M. G. Kendall, *The Advanced Theory of Statistics* (C. Griffin and Company, London, 1952), fifth edition, p. 211.

[18] H. Margenau and G. M. Murphy, *The Mathematics of Physics and Chemistry* (D. Van Nostrand Company, Inc., New York, 1953), p. 498.

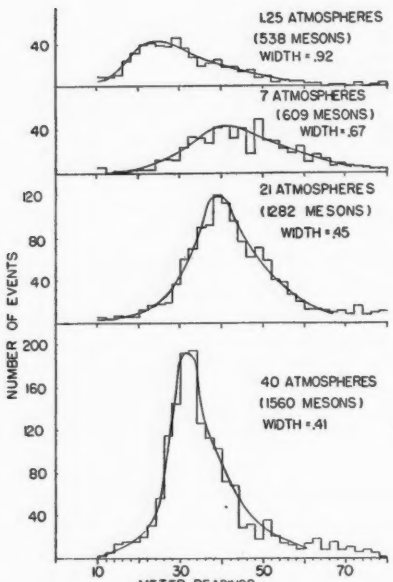


FIG. 4. Combined histograms for entire meson spectrum for various gas pressures in the high-pressure proportional counter.

quantile method are given by the following expressions:

$$q(a)=0.10, \quad q(b)=0.315, \quad q(c)=0.70,$$

for $u=2.4b$, where a, b, and c are, respectively, the lower half height, most probable, and upper half height value of the abscissa.

These values are very close to those obtained from the truncated theoretical curves,[12] with the exception of the lower quantile point. Here the theory indicates a value $q(a)=0.05$. Since it is just in this region that previous workers[19] have reported the greatest discrepancy between theory and experiment we have chosen to work with the experimentally determined best value of the lower quantile point.

The application of the quantile method is exceedingly simple and rapid, several successive approximations generally being sufficient to locate all points. We have subjected our procedure to the following thorough checks. With the data of Run 5, the effect upon the observed relativistic rise, of defining the quantile value of the most probable pulse height at various points within the range from 0.25 to 0.375 was studied. It was found that this rise was unaffected, within the limits of statistical accuracy of our data, over this entire range.

With the data of Run 1, in which the largest numbers of mesons were observed under any given set of conditions, an extensive comparison was made of the two methods. The results of this study, which are shown in Table III, indicate the accuracy with which the two methods check one another.

A further check was made with a synthetic distribution, constructed to be almost identical with the average shape of all our distributions, with a half-width of 0.71 and a quantile value of 0.315 for the most probable value of the abscissa. A set of 100 histograms, each containing 100 events, were obtained from this distribution by a random sampling method and the distributions in the observed values of the most probable and the half-width were obtained. The results achieved, which again were well within the standard deviation, indicated that for samples of the order of 100 events or greater the quantile method leads quickly to the same results as Beherens method and that the complicated expression previously given for the variance of the width at half-maximum is correct.

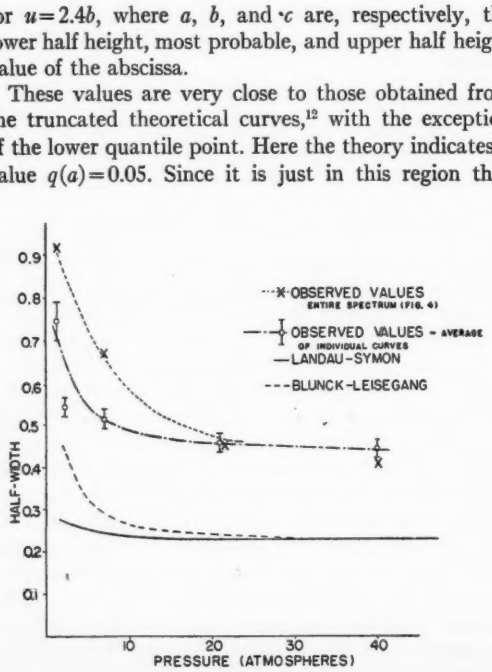


FIG. 5. Comparison of the theoretical and observed values of the width at half-maximum of the Landau distribution as a function of gas pressure.

V. RESULTS AND CONCLUSIONS

In Fig. 4 are shown four histograms for the pulse height distribution of the entire meson spectrum at four different pressures. Since each of these histograms is composed of many Landau distributions, shifted with respect to one another, no accurate curve fitting

TABLE III. Comparison of the results of the Beherens and quantile methods of analysis for the low-pressure proportional-counter data.

(Run 1) Group	No. of mesons in truncated histogram	Most probable pulse height (Arbitrary scale) Beherens	Most probable pulse height (Arbitrary scale) Quantile	Half-width Beherens	Half-width Quantile
A	463	25.5±0.5	26.0±0.5	0.55 ±0.05	0.54±0.06
B	503	28.8±0.5	28.5±0.5	0.515±0.04	0.53±0.05
C	950	32.0±0.5	32.3±0.4	0.50 ±0.03	0.52±0.04
D	591	34.0±0.5	35.0±0.5	0.48 ±0.04	0.53±0.05
		Weighted average		0.51 ±0.02	0.53±0.025

[19] P. Rothwell, Proc. Phys. Soc. (London) **B64**, 911 (1951).

procedure has been attempted, and the curves drawn in each case are simply the average of the most probable hand drawn curves as judged by nine independent observers. In this crude manner we have estimated the half-widths shown with each histogram in Fig. 4. Although the strong pressure dependence of the entire meson pulse height spectrum is quite apparent, the nature of the dependence does not become clear until the features of the histograms of the different energy bands are studied and the separate behaviors of the Landau distribution and of the density effect are determined.

In Fig. 5 the percentage widths at half-maximum are plotted against pressure for all our data along with the Landau-Symon and the Blunck-Leisegang theoretical widths. While the general shape of the experimental curve agrees well with the Blunck-Leisegang curve, the two appear to differ by a relatively constant amount even at our highest pressures. We know of no experimental defect in our work which would lead to an error of this type and magnitude and, as far as we are able to ascertain, all published results exhibit this tendency. This is demonstrated in Fig. 6 where we have plotted all such data available to us, the abscissa now being expressed as the logarithm of the amount of material traversed. While these data are for several gases for different experimental arrangements and for electrons as well as mesons, it is apparent that the prediction of decrease of width with increase of pressure is confirmed, but that a strong disagreement between the theoretical and observed magnitudes of the widths still exists.

The behavior of the density effect is exhibited in Fig. 7 where the most probable energy loss is plotted as a function of the mean momentum of the meson group for various values of gas pressure. The solid lines

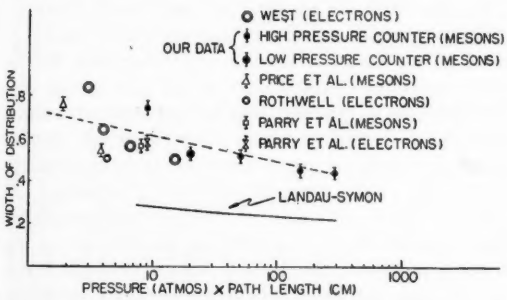


FIG. 6. Comparison of theoretical and all observed values of the width at half-maximum of the Landau distribution as a function of the logarithm of the amount of material traversed.

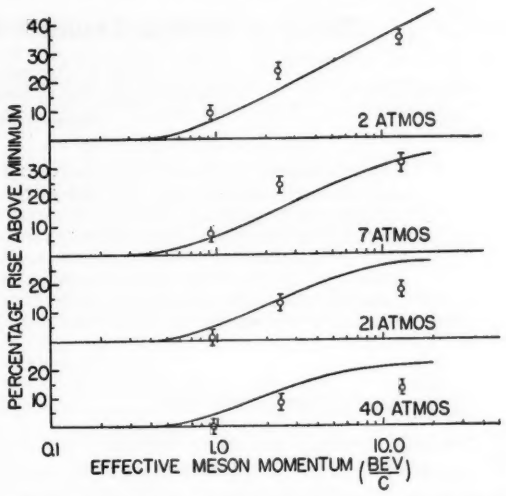


FIG. 7. Comparison of the theoretical rise of the most probable ionization energy loss (based on Sternheimer's calculations for the density effect) with the observed value at various gas pressures.

are the density-corrected curves as obtained from Sternheimer's[20] calculations. Within the limit of accuracy of our results the agreement is excellent; there may be some slight tendency towards a greater density effect than expected but we no longer consider this difference significant, contrary to a preliminary report emanating from this laboratory.[21]

Finally, it should be noted that both of the above pressure effects are required to explain the very strong pressure dependence of the pulse height distribution for the entire meson spectrum (top curve of Fig. 5 and all curves of Fig. 4) and that the rapidity with which the top curves of Fig. 5 converge indicates that the Fermi plateau is reached at relatively low gas pressures.

In conclusion the authors wish to express their thanks to Mr. Nestore Di Costanzo for his invaluable assistance in the construction of the high-pressure proportional counter, to Dr. W. A. Bowers for his helpful advice in the interpretation of the results and to Dr. G. E. Nicholson, Jr. and Dr. W. Hoeffding of the Department of Statistics for their invaluable aid in the methods of the treatment of the data. Finally, we wish to thank the Oak Ridge National Laboratories of the U. S. Atomic Energy Commission for the loan of the large quantity of lead required to perform the experiment.

[20] R. M. Sternheimer, Phys. Rev. **88**, 851 (1952).
[21] Palmatier, Meers, and Askey, Phys. Rev. **94**, 766(A) (1954).

Theory of Multiple Coulomb Scattering from Extended Nuclei*

LEON N. COOPER† AND JAMES RAINWATER
Department of Physics, Columbia University, New York, New York
(Received September 30, 1954)

Two independent methods are described for calculating the multiple scattering distribution for projected angle scattering resulting when very high-energy charged particles traverse a thick scatterer. The single scattering law for projected angle scattering is taken to be the Rutherford scattering law for projected angle scattering modified at small angles by electron shielding and at large angles by a nuclear form factor $\mathfrak{F}_n(\varphi/\varphi_0)$ which gives the effect of the finite nuclear size. The calculations can be carried through for any reasonable choice of $\mathfrak{F}_n$ and have been carried through for two suggested choices of $\mathfrak{F}_n$ for the examples of fast μ-meson scattering in lead slabs of thickness 2 cm and 5 cm, with good agreement for the two methods of calculation. The results are compared with the theories of Molière and Olbert.

1. INTRODUCTION

SEVERAL multiple scattering theories have been published[1-4] which are concerned with the angular distribution of particles passing, with no substantial loss of energy, through a thick material for which the single scattering law is the Rutherford cross section modified at small angles due to electron screening.

In this paper we shall present two methods by which a multiple scattering distribution for projected angle scattering can be obtained from more general single scattering laws than those already treated. We shall be concerned in particular with single scattering which is of Rutherford type for a wide middle region, but is modified both at large as well as at small angles. In scattering from atoms the modification at small angles is due to electron screening, while at large angles there can be modifications due to additional forces (as in the scattering of π mesons from carbon) or due to the finite extension of the nucleus and the subsequent deviations from Coulomb's law inside the nucleus. We consider only the case where relatively small φ are important so the approximation $\varphi \approx \sin\varphi \approx \tan\varphi$ is appropriate.

There has been a good deal of experimental interest recently[5,6] in the multiple scattering of μ mesons. Experimental results, after being compared with the results of the Molière[2] and Olbert[3] multiple scattering theories, have been interpreted as indicating the existence of an anomalous μ-meson-nuclear interaction. Among the many difficulties arising in the interpretation of these experiments (see Appendix B), one of the most obvious seemed to be the absence of a reliable estimate of the expected multiple scattering distribution from extended nuclei.

In the Molière multiple scattering theory the nucleus is treated as a point charge. The single scattering cross section is taken to be the Rutherford cross section modified, at small angles only, due to electron screening. This gives more scattering for large angles than would be expected from extended nuclei. In the Olbert theory an attempt is made to estimate the effect of the nuclear extension by multiplying the Molière projected angle single scattering law by a step function which cuts off all scattering beyond a given projected angle $\varphi_0 \approx \lambda/R$. This, however, gives a very great underestimate of the multiple scattering for angles larger than φ_0, as for large angles Olbert's distribution falls off as a Gaussian which soon is much smaller than even the coherent part of the single scattering law.

In what follows we shall outline two distinct procedures by which a multiple scattering distribution can be obtained from given single scattering laws. We deal with projected angle scattering as this is the usual experimental parameter. Our second method could be extended to include the total scattering angle. Sections 2 and 3 describe the two methods we have developed for dealing with multiple scattering problems. The results of these two sections have been used to calculate the expected multiple scattering distribution of relativistic μ mesons ($cp=1$ Bev) passing through 2 and 5 cm of lead. Agreement between the two methods is excellent. Appendix A contains a discussion of the single scattering cross section used for the above calculation. Appendix B gives a review of the experimental situation with regard to μ-meson scattering.

SECTION 2. DESCRIPTION OF THE FIRST METHOD OF COMPUTING $M(\varphi)$

The first method to be described consists of folding together several simpler distributions in a manner somewhat analogous to the actual effect of successive layers

* This research was supported in part by the joint program of the U. S. Office of Naval Research and the U. S. Atomic Energy Commission.

† Now at the Institute for Advanced Study, Princeton, New Jersey.

[1] E. J. Williams, Proc. Roy. Soc. (London) **A169**, 531 (1939); S. Goudsmit and J. L. Saunderson, Phys. Rev. **58**, 36 (1940); H. S. Snyder and W. T. Scott, Phys. Rev. **76**, 220 (1949); H. W. Lewis, Phys. Rev. **78**, 526 (1950).

[2] G. Molière, Z. Naturforsh. **2a**, 133 (1947); **3a**, 78 (1948).

[3] S. Olbert, Phys. Rev. **87**, 319 (1952); Annis, Bridge, and Olbert, Phys. Rev. **89**, 1216 (1953).

[4] H. A. Bethe, Phys. Rev. **89**, 1256 (1953).

[5] E. Amaldi and G. Fidecaro, Nuovo cimento **7**, 535 (1950); W. L. Whittemore and R. P. Shutt, Phys. Rev. **88**, 1312 (1952); George, Redding, and Trent, Proc. Phys. Soc. (London) **A66**, 533 (1953); B. Leontic and A. W. Wolfendale, Phil. Mag. **44**, 1091 (1953).

[6] M. L. T. Kanangara and G. S. Shrikantia, Phil. Mag. **44**, 1091 (1953).

of the scatterer. Although this method is simple in principle, it seemed on first inspection that it would be extremely tedious to carry through such folding operations. This did not prove to be the case when the computing techniques described below were used, and we give a rather detailed description of the procedure for this reason.

The method is based on the observation that if $f(\varphi)d\varphi$, the probability of a single scattering through φ to $\varphi+d\varphi$ to one side in projected angle, is given, then the multiple scattering distribution $M(\varphi)$ is completely defined. [Note that $f(\varphi)$ is not normalized to unity, but integration over all angles gives the average number of single scatterings in transversing the sample.] Furthermore, if the actual scattering slab were replaced by a series of consecutive slabs A, B, C having single scattering laws $f_A(\varphi)$, $f_B(\varphi)$, $\cdots$, where

$$f(\varphi)=f_A(\varphi)+f_B(\varphi)+\cdots, \tag{1}$$

then the same multiple scattering distribution results on traversing all of the slabs in series. If $M_A(\varphi)$, $M_B(\varphi)$, $\cdots$ are the separate multiple scattering distributions for A, B, $\cdots$, then $M(\varphi)$ results on folding $M_A, M_B, \cdots$ together.

For the case of two components,

$$M(\varphi)=\int_{-\infty}^{\infty} M_A(\varphi_1)M_B(\varphi-\varphi_1)d\varphi_1. \tag{2}$$

We require that only small φ are important so $\sin\varphi \approx \tan\varphi \approx \varphi$, and effectively consider φ as ranging from $-\infty$ to $+\infty$.

The distribution law for small angles is just the Rutherford scattering law modified because of electron shielding. Following Molière, we represent this as

$$f'(\varphi)=\tfrac{1}{2}Q(\varphi^2+\varphi_m^2)^{-1}, \tag{3}$$

where φ is the projected angle, φ_m is the screening angle,

$$\varphi_m=\frac{1.14m_ec^2Z^{\frac{1}{3}}}{137cp}\left[1.13+3.76\left(\frac{Z}{137\beta}\right)^2\right]^{\frac{1}{2}}\approx\lambda\!\!\!^{-}/a,$$

and $Q=4\pi(Nt/A)(Ze^2/pv)^2$. Here (Nt/A) gives the scatterer thickness in atoms/cm^2; p is the momentum of the incoming particle; $\lambda\!\!\!^{-}=\hbar/p$; v is the incoming velocity; a is the "Fermi-Thomas radius" of the atom $=1.67\times10^4Z^{-\frac{1}{3}}(e^2/m_ec^2)$. Z is the atomic number of the scattering material (the incident particle is taken to be singly charged) and m_e is the electron mass.

The modification in the above distribution law at larger angles is given by multiplying $f'(\varphi)$ above by $\mathfrak{F}_N(\varphi/\varphi_0)$ to give

$$f(\varphi)=\tfrac{1}{2}Q(\varphi^2+\varphi_m^2)^{-1}\mathfrak{F}_N(\varphi/\varphi_0). \tag{4}$$

$\mathfrak{F}_N(\varphi/\varphi_0)$ is discussed in Appendix A.

The trick of this method consists of selecting some angle φ' which is a little smaller than the rms width of the Gaussian which approximates $M(\varphi)$ at small angles. $f_A(\varphi)$ is then set equal to $f(\varphi)$ for $|\varphi|<\varphi'$ and $=0$ for $|\varphi|\geqslant\varphi'$, while $f_B(\varphi)=0$ for $|\varphi|<\varphi'$ and $=f(\varphi)$ for $|\varphi|\geqslant\varphi'$. The multiple scattering due to f_A is given with good accuracy by the Gaussian:

$$M_A(\varphi)=(2\pi\langle\varphi^2\rangle)^{-\frac{1}{2}}\exp(-\varphi^2/2\langle\varphi^2\rangle), \tag{5}$$

where

$$\langle\varphi^2\rangle=Q\int_0^{\varphi'}\varphi^3(\varphi^2+\varphi_m^2)^{-2}d\varphi, \tag{6}$$

since $\mathfrak{F}_N(\varphi/\varphi_0)\approx1$ for $|\varphi|\leqslant\varphi'$. Also $\varphi'\gg\varphi_m$ for a reasonably thick scatterer, so we can simplify the resulting expression:

$$\langle\varphi^2\rangle=Q\left[\frac{-x}{(1+x^2)^{\frac{1}{2}}}\right]_0^{\varphi'/\varphi_m}+\left[\ln(x+(1+x^2)^{\frac{1}{2}})\right]_0^{\varphi'/\varphi_m},$$

so

$$\langle\varphi^2\rangle\approx Q\{\ln(2\varphi'/\varphi_m)-1\}. \tag{7}$$

It is convenient to use the parameter $y=\varphi/\varphi_0$ and replace $f(\varphi)$, Q, φ_m, and φ' by the equivalent quantities $g(y)$, B, y_m, and y'. For the case of a μ meson with $cp=1$ Bev and a 2-cm lead scatterer, $\varphi_0=0.0304$ radians$=1.74°$, $B=0.126$, and $y_m=9.6\times10^{-4}$. For the choice $y'=0.5$ we obtain $\langle y^2\rangle=0.75$, showing that this is a satisfactory choice for y'. Then, for this case,

$$M_{A1}(y)=(1.50\pi)^{-\frac{1}{2}}\exp(-y^2/1.50), \tag{8}$$

where the subscript 1 on M_{A1} emphasizes that this method gives the distribution in terms of $y=\varphi/\varphi_0$. (The distribution function $M_2(x)$ of the next section is given in terms of the angle characteristic of the Molière theory.)

The multiple scattering distribution $M_{B1}(y)$ corresponding to single scattering angles $\geqslant\varphi'=y'\varphi_0$ is obtained using the following considerations. If we chose, not the full scatterer thickness, but some sufficiently small fraction α, then the single scattering law $\alpha g(y)$ for $|y|>y'$ would have essentially unit weighting for no scatterings at all, and very small weightings for all $|y|>y'$. The corresponding multiple scattering distribution would then be identical to the single scattering law for $|y|>y'$ since there is negligible likelihood of more than one scattering. If this distribution is folded together with itself the multiple scattering is obtained for fraction 2α of the total thickness. This distribution differs from the single scattering law by double scattering terms proportional to α^2, so the *fractional* deviation is proportional to $\alpha\ll1$. If the multiple scattering law for thickness 2α is folded together with itself the multiple scattering law for thickness 4α results. For $|y|>y'$ this distribution has a *fractional* deviation from double that for 2α of an amount double that of the previous process (always neglecting higher order effects). Thus the net fractional correction from the single scattering law is 3 times that obtained in the first folding. Subsequent foldings for 8α, 16α, 32α, etc. have *net* fractional corrections (1+2+4), (1+2+4+8),

(1+2+4+8+16), etc., times that of the folding starting with thickness α. These sums, 7, 15, 31, $\cdots$, would be 8, 16, 32, $\cdots$, if we had started our folding process with an infinitesimal fraction of α, since going from $\alpha/2$ to α gives $\frac{1}{2}$, from $\alpha/4$ to $\alpha/2$ gives $\frac{1}{4}$, etc., for the extra series terms $(\frac{1}{2}+\frac{1}{4}+\cdots)=1$. We thus note that the net fractional correction after any stage of folding (neglecting higher order effects) is just double the fractional correction obtained by the given step (i.e., comparing the multiple scattering for $y\neq 0$ for the given thickness with double that for half thickness). We make use of the above feature to select as our starting thickness a fraction 2^{-n} of the total for which the fraction correction of any of the points in the first folding is not too large. This fractional correction is then doubled to account with good accuracy for the effect of not starting the process with an infinitely thin sample. In the calculations for 2-cm lead and $cp=1$ Bev it was found to be suitable to start with $\frac{1}{8}$ the sample thickness, while $\frac{1}{16}$ the sample thickness was suitable for 5-cm lead. The actual folding operations can be carried through with good accuracy by replacing the continuous $f_B(\varphi)$ by a discontinuous function having values only at regular grid points. For the 2-cm lead case where $y'=0.5$, the region 0.5 to 0.7 was represented by $0.2\,g(0.6)$, etc. so values were defined only at $y_j=0$, ± 0.6, ± 0.8, ± 1.0, $\cdots$, with a maximum $y_j=6.0$ used in the calculations. For the 5-cm lead case we used $y'=1.2$ and 0.4 intervals to $y=10.0$. It was found convenient and permissible in these cases to neglect contributions to values of $y\neq 0$ where $|y|<y'$. (In cases where these contributions are not negligible it requires only slight additional computing time to include them.) The remaining computing tricks can best be illustrated in terms of the example of the calculation for the 2-cm lead case where we started with $\frac{1}{8}$ of the total thickness. Let $l_{1/8}{}^0(y_j)$ be the lumped weights given to the grid points 0, ± 0.6, ± 0.8, $\cdots$, ± 6.0 to represent the single scattering law $(0.2/8)g(y_j)$ for $|y_j|>0.5$, with

$$l_{1/8}{}^0(0)=1-2\sum_{y_j>0} l_{1/8}{}^0(y_j). \tag{9}$$

A calculation sheet is now prepared in which values of $y_j=0, 0.2, 0.4, \cdots, 6.0$ are entered in the first column. The values of $l_{1/8}{}^0(y_j)$ are entered in the second column. An uncorrected multiple scattering distribution $l_{1/4}{}^0(y)$ is then computed, where

$$l_{1/4}{}^0(y_j)=\sum_k l_{1/8}{}^0(y_k)l_{1/8}{}^0(y_j-y_k). \tag{10}$$

Use is made of the fact that terms for $y_k\neq y_j/2$ appear twice and we can write

$$l_{1/4}{}^0(y_j)=2\sum_{y_k<y_j/2} l_{1/8}{}^0(y_k)l_{1/8}{}^0(y_j-y_k) +[l_{1/8}{}^0(y_j/2)]^2. \tag{11}$$

The calculations thus proceed as follows. In the third column the values of $l_{1/8}{}^0(y_j)$ are multiplied by $l_{1/8}{}^0(0)$. The first number appearing in the column ($y_j=0$ here) is circled and is only counted once in the subsequent summing. The next column entry starts at $y_j=1.2$ and contains products $l_{1/8}{}^0(0.6)\,l_{1/8}{}^0(y_j-0.6)$. Again the first term at $y_j=1.2$ is circled. The next column starts at $y_j=1.6$ and contains terms $l_{1/8}{}^0(0.8)\,l_{1/8}{}^0(y_j-0.8)$. Subsequent columns are formed similarly to produce a triangular array (requiring less than 1 hour of slide rule computing time). Contributions from negative y_k are usually negligible but the terms for the first few negative values of y_k can be added if necessary. The rows are then added as indicated by Eq. (11) to give $l_{1/4}{}^0(y_j)$. The function $l_{1/4}{}^0(y_j)$ is then compared with $2l_{1/8}{}^0(y_j)$ for $y_j\neq 0$ and the differences are added to $l_{1/4}{}^0(y_j)$ to obtain the corrected multiple scattering distribution $l_{1/4}(y_j)$. This is similarly folded with itself to give $l_{1/2}(y_j)$, and a repetition gives $l(y_j)$ which corresponds to $M_{B1}(y)$. Values of $M_{A1}(y)$ [Eq. (8)] are computed for $y_j=0, 0.2, 0.4, \cdots, 6.0$. The final $M_1(y)$ is just the sum of Gaussians centered at 0, ± 0.6, ± 0.8, $\cdots$ ± 6.0, $\cdots$.

$$M_1(y_j)=\sum_{y_k} l(y_k)M_{A1}(y_j-y_k). \tag{12}$$

The value of $M_1(y_j)$ for any y_j is computed using Eq. (12). For y_j near 6.0 the contributions from $y_k>6.0$ must be estimated. Fortunately the single scattering $g(y)$ and $l(y_k)$ are decreasing positive functions of y, so the terms in Eq. (12) have a maximum for $y_k<y_j$ and the remainder can be estimated by noting the behavior of the series for smaller values of y_j.

The above observations suggest an easy method for estimating $M_1(y)$ for large y in terms of $M_{A1}(y)$ and $M_{B1}(y)$ without actually carrying through the process indicated by Eq. (12). If $M_{B1}(y)$ can be approximated by a straight line on a semilog plot of the form

$$M_{B1}(y_k+x)=M_{B1}(y_k)e^{-x/a}, \tag{13}$$

then Eq. (2) gives

$$M_1(y_j)=M_{B1}(y_j)\exp(\langle y^2\rangle/2a^2). \tag{14}$$

This may be regarded as increasing $M_{B1}(y_j)$ by a certain amount on the semilog plot to produce $M_1(y_j)$. Alternatively, it can be written

$$M_{B1}(y_j)=M_1(y_j+\Delta) \tag{15}$$

corresponding to a sidewise displacement Δ of $M_{B1}(y_j)$ to generate $M_1(y_j+\Delta)$. The quantity Δ is obtained by plotting both $M_{A1}(y)$ and $M_{B1}(y)$ on the same semilog plot as in Fig. 1. Let Δ_j' be the value of y at which $M_{A1}(y)$ has the same (logarithmic) slope that $M_{B1}(y)$ has at $y=y_j$. Using Eq. (13) shows that $\Delta'=\langle y^2\rangle/a$, so $\Delta=\Delta'/2$. Here a is assumed to be a slowly varying function of y. (The plot by comparison can be made rapidly, using two straight edges.) When the semilog plot of $M_{B1}(y)$ also has curvature so

$$M_{B1}(y_k+x)\approx M_{B1}(y_k)\,e^{-x/a}\exp(x^2/2b^2), \tag{16}$$

then the equation analogous to Eq. (14) is

$$M_1(y_j)\approx M_{B1}(y_j)[b^2/(b^2-\langle y^2\rangle)]^{\frac{1}{2}}\exp\left(\frac{-b^2\langle y^2\rangle}{2a^2(b^2-\langle y^2\rangle)}\right),$$

which, for $b^2\gg\langle y^2\rangle$, can be written approximately as

$$M_1(y_j)\approx M_{B1}(y_j)\exp\left(\frac{\langle y^2\rangle}{2a^2}[1+(a^2+\langle y^2\rangle)/b^2]\right). \quad (17)$$

Inspection of the final result shows that this corresponds again to a sidewise displacement of M_{B1} by $\Delta\approx 0.6\Delta'$ to generate M_1 in the examples considered. An alternate method of sidewise displacement is obtained by noting that $M_{A1}(x)M_{B1}(y-x)$ has its maximum at Δ', about which it resembles a displaced Gaussian. This suggests that $M_1(y_j)$ is generated mainly from $M_{B1}(y_k)$ in the region $y_k\approx y_j-\Delta'$ so $M_1(y)$ should be generated using $M_{B1}(y-\Delta')$ multiplied by $\exp(-\Delta'^2/2\langle y^2\rangle)$. [The last factor is the ratio $M_{A1}(\Delta')/M_{A1}(0)$.] These methods of generating the approximate curve for $M_1(y)$ can be carried out rapidly

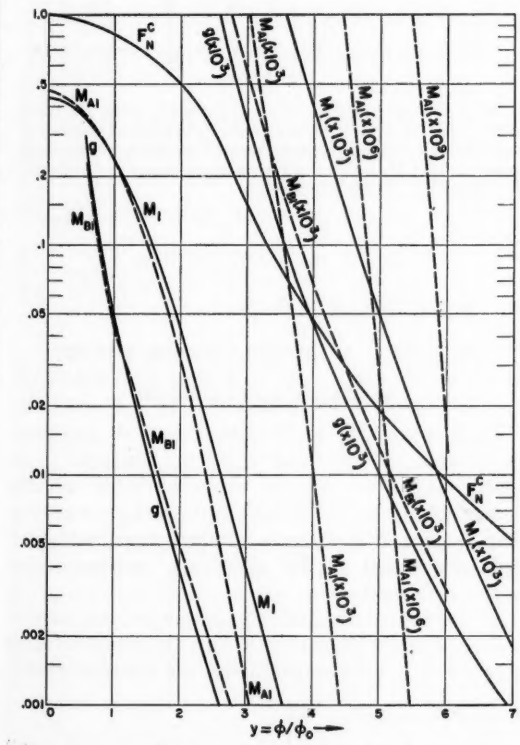


FIG. 1. Curves illustrating the first method of computing the multiple scattering distribution $M_1(y)$ for $cp=1$ Bev in 2 cm Pb, including only coherent scattering. g=single scattering law; M_{A1}=multiple scattering Gaussian for single scatterings through $y\lesssim 0.5$; M_{B1}=multiple scattering distribution for single scatterings through $y>0.5$; $F_N{}^C$ is the assumed coherent nuclear form factor. Multiply y by 1.74 for Bev degrees. A nuclear size $R=1.0A^{\frac{1}{3}}\times10^{-13}$ cm was used for all of the examples of the figures, as discussed in Appendix A.

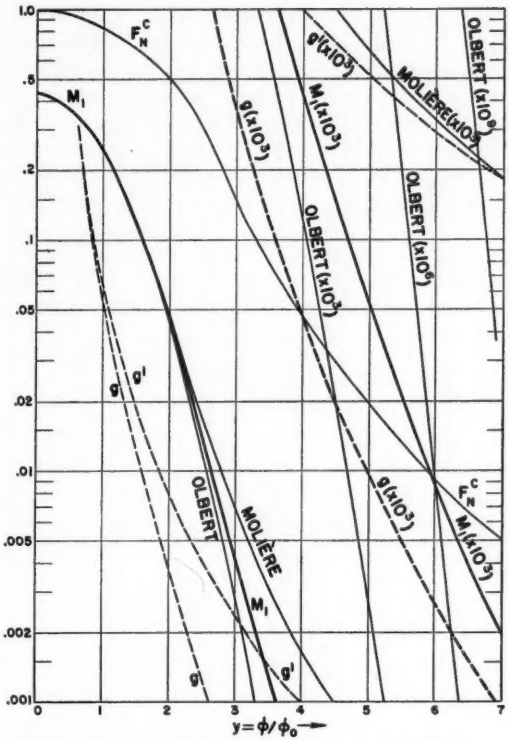


FIG. 2. Curves for the same case as in Fig. 1. The Molière and Olbert multiple-scattering distributions are shown for comparison with $M_1(y)$. The Olbert distribution is for a single-scattering cutoff angle of $y=1.0$. The point-nucleus and extended-nucleus single-scattering distributions are g' and g.

by simple displacements on the semilog plot and are quite instructive in giving insight into the behavior of $M_1(y)$. The difference between $M_{B1}(y)$ and $g_B(y)$ decreases rapidly if the choice of the dividing angle y' is increased. This is compensated largely by an increase in the width of the Gaussian $M_{A1}(y)$, and thus in the required sidewise displacement Δ' to generate $M_1(y)$ from $M_{B1}(y)$. It is of interest that the second method (of the next section), at large angles, just folds *the single scattering law* at large angles with a Gaussian characteristic of the Molière theory, which is somewhat wider than $M_{A1}(y)$ for the method of selecting y' discussed above.

Figures 1, 2, and 3 show the characteristic functions when only elastic processes are included. It is seen that $M_1(y)$ remains considerably greater than g_B in a region where the Molière curve is almost the same as the point nucleus single scattering curve. This is consistent with the above discussion since the logarithmic slope of the point nucleus $g(y)$ is much smaller than the extended nucleus $g(y)$ so the required sidewise displacement is less, and the result of a given sidewise displacement is also less [using Eq. (15)]. When inelastic scattering is included, Figs. 4 and 5, the curves approach those for a point nucleus, decreased by a factor Z^{-1} at very large y.

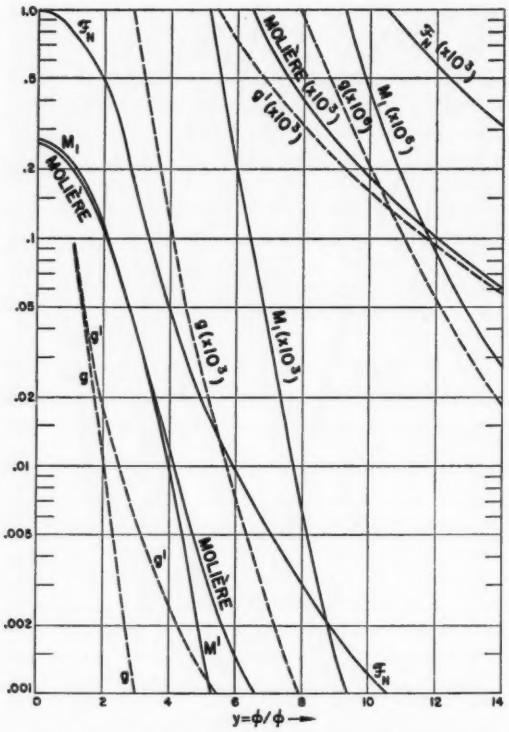


FIG. 3. Curves for $cp=1$ Bev and 5 cm Pb, using only coherent scattering. g^1, g, $\mathfrak{F}_N$, and M_1 have the same meaning as in the preceding figures. Multiply y by 1.74 for Bev degrees.

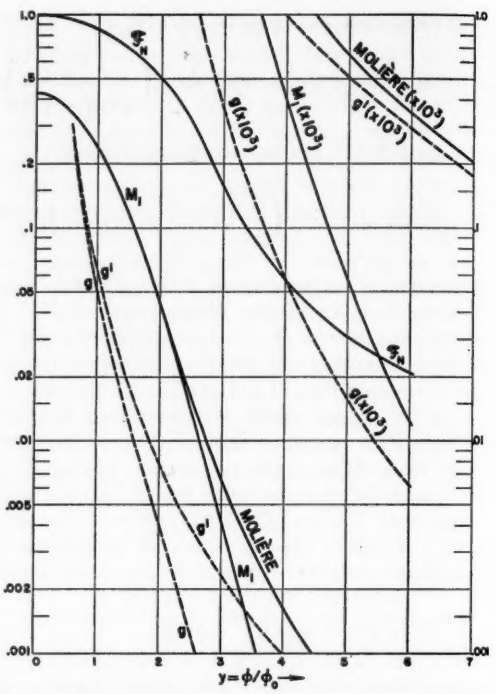


FIG. 4. Curves for $cp=1$ Bev and 2 cm Pb, using $\mathfrak{F}_N$ for the total elastic plus inelastic scattering. g'=point-nucleus single-scattering law; $g=g'\mathfrak{F}_N$ is the assumed extended-nucleus single-scattering law; M_1 is the resulting multiple-scattering distribution. Multiply y by 1.74 for Bev degrees.

It is instructive to make a further comparison of this method and the second method (of the next section) for $y>4$, say. This method folds together the small-angle multiple scattering Gaussian with the large angle *multiple* scattering distribution, while the second method essentially folds together a *wider* Gaussian and the law for *single* scattering. The second method uses an expansion in terms of the parameter $(4G)^{-1}$ described below, neglecting terms of order $(4G)^{-2}$ and beyond. This very roughly corresponds to neglecting multiple large-angle scattering, so the first method would be expected to be more reliable in cases where such effects are not negligible.

As a final point of interest, we note that the curves given in terms of y apply for all relativistic momenta where $\beta\approx 1$. In the example considered, y is in units of 1.74 Bev degrees. Thus multiplying the abscissa, and dividing the ordinate by 1.74 gives the distribution per Bev degree unit of $(cp\varphi)$. This is shown by noting that $f(\varphi)d\varphi=g(y)dy$, so

$$g(y)=\tfrac{1}{2}B(y^2+y_m^2)^{-\frac{1}{2}}\mathfrak{F}_N(y),$$

where

$$B=Q\varphi_0^{-2}=4\pi(ZR/137\beta)^2(Nt/A),$$

$$y_m=\varphi_m/\varphi_0=\frac{1.14R}{137^2r_{0e}}[1.13+3.76(Z/137\beta)^2]^{\frac{1}{2}},$$

and

$$r_{0e}=e^2/m_ec^2\approx 2.82\times 10^{-13}\text{ cm}$$

are independent of p for $\beta\approx 1$.

SECTION 3. MODIFIED MOLIÈRE THEORY

A. Methods and Notation

To introduce the mathematical methods and notation used in this section we review the derivation of the general expression for the projected angle multiple scattering distribution for an arbitrary single scattering cross section. The derivation follows that of Molière[2] and Olbert[3] and applies when only relatively small angles are important.

If $f(\varphi)d\varphi$ is the probability that an incident particle undergo a single scattering through the projected angle φ to $\varphi+d\varphi$ in its passage through the scatterer, then

$$P(\varphi_1,\varphi_2,\cdots\varphi_n)d\varphi_1\cdots d\varphi_n$$
$$=e^{-\sigma_t}f(\varphi_1)f(\varphi_2)\cdots f(\varphi_n)d\varphi_1d\varphi_2\cdots d\varphi_n, \quad (18)$$

where

$$\sigma_t\equiv\int_{-\infty}^{\infty}f(\varphi)d\varphi$$

is the probability that the particle in passing through the material has exactly n single scatterings through

the projected angles $\varphi_1, \varphi_2, \cdots, \varphi_n$ with angular intervals $d\varphi_1, d\varphi_2, \cdots, d\varphi_n$. $e^{-\sigma_t}$ is the probability that there be no other scatterings besides $\varphi_1, \varphi_2 \cdots \varphi_n$. Since we are interested only in the final angle φ we integrate over intermediate angles and put

$$P^{(n)}(\varphi)d\varphi=\frac{d\varphi}{n!}e^{-\sigma_t}\int_{-\infty}^{\infty}\cdots\int f(\varphi_1)f(\varphi_2)\cdots$$

$$\times f(\varphi-\varphi_1-\varphi_2-\cdots-\varphi_{n-1})d\varphi_1\cdots d\varphi_{n-1}, \quad (19)$$

where $P^{(n)}(\varphi)d\varphi$ is the probability that the particle emerge at a projected angle between φ and $\varphi+d\varphi$ after exactly n scatterings in the material. The factor $1/n!$ occurs because in the integration over intermediate angles the $n!$ permutations of $\varphi_1\cdots\varphi_n$ have each been counted. Since Eq. (18) holds for independent events the order of the n scatterings is irrelevant.

Now introduce the Fourier representation of the single scattering cross section. Let

$$f(\varphi)=\frac{1}{2\pi}\int_{-\infty}^{\infty}g(\xi)e^{i\xi\varphi}d\xi, \quad (20)$$

$$g(\xi)=\int_{-\infty}^{\infty}f(\varphi)e^{-i\xi\varphi}d\varphi, \quad (21)$$

[where $g(\xi)$ is not related to $g(y)$ of the previous section]. Putting Eq. (20) into Eq. (19) gives

$$P^{(n)}(\varphi)d\varphi=\frac{d\varphi}{n!}e^{-\sigma_t}(2\pi)^{-n}\int_{-\infty}^{\infty}\cdots\int g(\xi_1)\cdots$$

$$\times g(\xi_n)e^{i\xi_1\varphi_1}\cdots e^{i\xi_n(\varphi-\varphi_1-\cdots-\varphi_{n-1})}d\xi_1\cdots$$

$$\times d\xi_n d\varphi_1\cdots d\varphi_{n-1}. \quad (22)$$

The integrations over $\varphi_1, \cdots \varphi_{n-1}$ give delta functions which, after the $\xi_1\cdots\xi_{n-1}$ integrations yield

$$P^{(n)}(\varphi)d\varphi=\frac{d\varphi}{2\pi}e^{-\sigma_t}\int_{-\infty}^{\infty}\frac{[g(\xi)]^n}{n!}e^{i\xi\varphi}d\xi. \quad (23)$$

The probability of a final projected angle φ after any number of scatterings is just the sum of the $P^{(n)}(\varphi)d\varphi$ over all values of n.

$$P(\varphi)d\varphi\equiv\sum_{n=0}^{\infty}P^{(n)}(\varphi)d\varphi=\frac{d\varphi}{2\pi}e^{-\sigma_t}\int_{-\infty}^{\infty}\sum_{n=0}^{\infty}\frac{[g(\xi)]^n}{n!}e^{i\xi\varphi}$$

$$=\frac{d\varphi}{2\pi}\int_{-\infty}^{\infty}e^{i\xi\varphi}e^{g(\xi)-g(0)}d\xi, \quad (24)$$

where we use the result from Eq. (21):

$$\sigma_t\equiv\int_{-\infty}^{\infty}f(\varphi)d\varphi=g(0).$$

From this it can be seen that the solution of the multiple-scattering problem is equivalent to the evaluation of the integral in Eq. (24). Molière[2] has given an evaluation of this integral in the case of a screened Coulomb field. The main trick in evaluating the integral lies in the observation that the existence of multiple scattering will smear out fine-grained irregularities in the final distribution so the high Fourier components give a negligible contribution.

Molière used the single-scattering law of Eq. (3) appropriate for a point nucleus with electron shielding. We are interested in the case where this is modified by the nuclear "form factor" $\mathfrak{F}_N(\varphi/\varphi_0)$ as in Eq. (4). From the discussion of Appendix A we note that

$$\mathfrak{F}_N(\varphi/\varphi_0)\equiv\begin{cases}1 & \text{for small } |\varphi/\varphi_0|\\ Z^{-1} & \text{for large } |\varphi/\varphi_0|\end{cases} \text{ (our choice)}, \quad (25)$$

with the rapid change occurring when $\varphi\sim 2\varphi_0$.

Olbert's method consisted of using a step function for $\mathfrak{F}_N$:

$$\mathfrak{F}_N(\varphi/\varphi_0)=\begin{cases}1 & \text{for } |\varphi|\leq\varphi_0\\ 0 & \text{for } |\varphi|>\varphi_0.\end{cases} \text{ (Olbert)} \quad (26)$$

As mentioned in the introduction, this gives a very large underestimate of the multiple scattering for

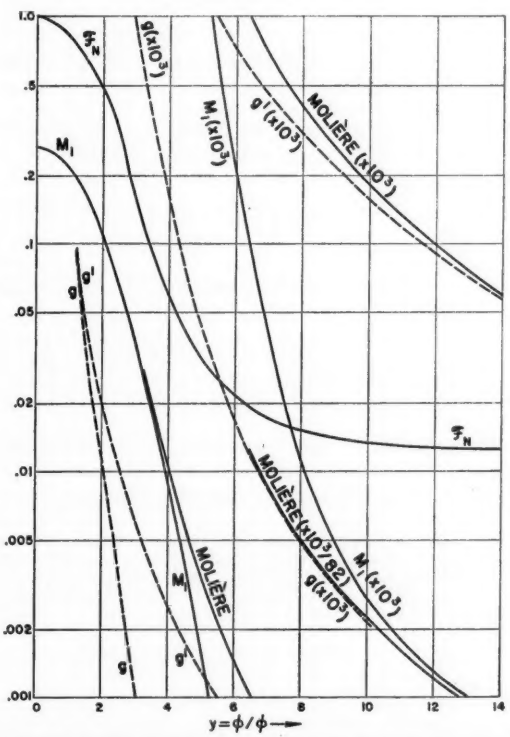


FIG. 5. Curves for $c\beta=1$ Bev and 5 cm Pb using the total elastic plus inelastic scattering. The symbols have the same meaning as in the preceding figures.

$\varphi \gg \varphi_0$ where the Olbert function falls off as a Gaussian while, for any reasonable form factor, the multiple-scattering distribution lies above the single-scattering distribution $f(\varphi)$ for large φ, and thus far above the Olbert distribution.

B. First Derivation of $M_2(x)$

We first treat the multiple scattering from extended nuclei as a correction to the Molière theory by setting $\mathfrak{F}_N=[1+(\mathfrak{F}_N-1)]$, where $(\mathfrak{F}_N-1)$ gives the correction term. This is the method used by Olbert for his step function. Such a treatment yields satisfactory results for small angles, but is inconvenient for large angles where a second method (described following this) is preferred. Equation (4) is written

$$f(\varphi)=\tfrac{1}{2}Q(\varphi^2+\varphi_m{}^2)^{-\frac{1}{2}}[1+(\mathfrak{F}_N-1)]. \tag{27}$$

Then

$$g(\xi)-g(0)=\frac{Q}{2}\int_{-\infty}^{\infty}\frac{(e^{-i\xi\varphi}-1)d\varphi}{(\varphi^2+\varphi_m{}^2)^{\frac{1}{2}}} +\frac{Q}{2}\int_{-\infty}^{\infty}\frac{(e^{-i\xi\varphi}-1)(\mathfrak{F}_N-1)d\varphi}{(\varphi^2+\varphi_m{}^2)^{\frac{1}{2}}}. \tag{28}$$

The first integral has been treated by Molière[2,7] and to a sufficient degree of accuracy, is

$$\frac{Q}{2}\int_{-\infty}^{\infty}\frac{(e^{-i\xi\varphi}-1)d\varphi}{(\varphi^2+\varphi_m{}^2)^{\frac{1}{2}}}\approx Q\left[\frac{\xi^2}{4}\ln\left(\frac{\gamma^2\varphi_m{}^2}{4e}\xi^2\right)\right]. \tag{29}$$

We now introduce the various parameters typical of the Molière theory, following the notation of Olbert:

$$G=-\tfrac{1}{2}\ln\left(\frac{\gamma^2\varphi_m{}^2}{e}\frac{1}{2G\varphi}\right),$$

$$x=(2GQ)^{-\frac{1}{2}}\varphi, \quad \eta=(2GQ)^{\frac{1}{2}}\xi; \tag{30}$$

and

$$G\approx 5.66+1.24\log_{10}\left[\frac{Z^{4/3}A^{-1}t}{1.13\beta^2+3.76(Z/137)^2}\right]$$

is independent of the momentum in the relativistic region.

Putting Eq. (29) in (28), and using (30) gives

$$s(\eta)\equiv g\left(\frac{\eta}{(2GQ)^{\frac{1}{2}}}\right)-g(0)=-\frac{\eta^2}{4}+\frac{1}{2G}\left[\frac{\eta^2}{4}\ln\frac{\eta^2}{4}+\frac{1}{2}I\right] \tag{31}$$

where

$$I=\int_{-\infty}^{\infty}\frac{(e^{-i\eta x'}-1)[\mathfrak{F}_N(x'/x_0)-1]dx'}{(x'^2+x_m{}^2)^{\frac{1}{2}}}.$$

Inserting this result into Eq. (24) we obtain in terms of the parameter x:

$$M_2(x)dx=\frac{dx}{2\pi}\int_{-\infty}^{\infty}d\eta e^{i\eta x}\exp\left[-\frac{\eta^2}{4}+\frac{1}{2G}\left(\frac{\eta^2}{4}+\frac{1}{2}I\right)\right]. \tag{32}$$

We use the fact that G is a reasonably large number (for moderate foil thickness G varies from about 7 to 15) and expand the portion of the exponential which is multiplied by $(2G)^{-1}$. This gives, to first order in $(2G)^{-1}$,

$$M_2(x)dx=\frac{dx}{2\pi}\int_{-\infty}^{\infty}d\eta\exp(i\eta x-\eta^2/4) \times\left[1+\frac{1}{2G}\left(\frac{\eta^2}{4}\ln\frac{\eta^2}{4}\right)+\frac{1}{4G}I+\cdots\right]. \tag{33}$$

After performing the integrations over η, this becomes

$$M_2(x)dx=\frac{\exp(-x^2)dx}{\sqrt{\pi}}+\frac{dx}{4G}[f'(x;\infty)-K(x)], \tag{34}$$

where

$$K(x)=\frac{1}{\sqrt{\pi}}\int_0^{\infty}\frac{dx'[1-\mathfrak{F}_N(x'/x_0)]}{(x'^2+x_m{}^2)^{\frac{1}{2}}}\{\exp[-(x+x')^2] +\exp[-(x-x')^2]-2\exp(-x^2)\}$$

and

$$f'(x;\infty)=\frac{1}{\pi}\int_{-\infty}^{\infty}d\eta\exp(i\eta x-\eta^2/4)\left(\frac{\eta^2}{4}\ln\frac{\eta^2}{4}\right).$$

$f'(x;\infty)$ is the Molière function. The integral $K(x)$ which occurs in Eq. (34) must be evaluated numerically for a given $\mathfrak{F}_N(x'/x_0)$. Equation (34) is seen to have the form that was desired originally. The integral which contains the effect of the nuclear extension occurs as a correction to $\exp(-x^2)/\sqrt{\pi}+(1/4G)f'(x;\infty)$, which is the Molière multiple scattering distribution for a point nucleus in terms of the projected angle parameter x. This corresponds to $\mathfrak{F}_N=1$ for all x. If $\mathfrak{F}_N(x/x_0)$ is the Olbert step function, Eq. (26), the correction term becomes the Olbert correction function $K(x;x_0)$ is use is made of the fact that $x_0 \gg x_m$.

Using Eq. (34) the multiple scattering distribution can be determined for a particular $\mathfrak{F}_N(\varphi/\varphi_0)$. However, Eq. (34) becomes inconvenient for large values of x, because in that case $K(x)$ becomes almost equal to $f'(x;\infty)$ and the difference between two large numbers must be used to give a small one. For large values of x therefore it is better to treat the modified cross section directly, rather than as a correction to the Molière distribution. This is done as below.

C. Second Derivation of $M_2(x)$

Consider again the Fourier transform of the single scattering law.

$$g(\xi)-g(0)=\frac{Q}{2}\int_{-\infty}^{\infty}\frac{(e^{-i\xi\varphi'}-1)\mathfrak{F}_N(\varphi'/\varphi_0)d\varphi'}{(\varphi'^2+\varphi_m{}^2)^{\frac{1}{2}}}. \tag{35}$$

[7] Also see H. A. Bethe (reference 4) for a discussion of this integral.

Introducing the Molière parameters x, η, and G as defined Eq. (30), we get

$$g\left(\frac{\eta}{(2GQ)^{\frac{1}{2}}}\right)-g(0)\equiv s(\eta)$$

$$=\frac{1}{2G}\int_0^\infty \frac{\mathfrak{F}_N(x'/x_0)(\cos\eta x'-1)dx'}{(x'^2+x_m^2)^{\frac{1}{2}}} \quad (36)$$

and

$$M_2(x)dx=\frac{dx}{2\pi}\int_{-\infty}^{\infty} e^{i\eta x}e^{s(\eta)}d\eta. \quad (36a)$$

Equation (36) of course is the Molière-Fourier transform if $\mathfrak{F}_N(x'/x_0)=1$. In the case of the point nucleus, this integral is evaluated through the observation that there exists an angle $x'=\kappa$ at which the integral can be split such that $\kappa\gg x_m$ but $\kappa\ll 1/\eta_c$, where η_c is the frequency in the neighborhood of the main Fourier components.[4] Then in the integration up to κ the factor $(\cos\eta x'-1)$ can be put equal to $-(\eta x')^2/2$ and, in the integration from κ to ∞, $(x'^2+x_m^2)^{-\frac{1}{2}}$ becomes x'^{-3}. Both integrations can be performed, and the splitting point κ cancels up to terms of the order κ^2, which are very small.

In the case with which we deal the argument is modified as follows. Again we split the integral at an angle κ where $x_m\ll x'=\kappa\ll\eta_c^{-1}$. Now we make use of the important property of the form factor given in Eq. (25) that $\mathfrak{F}_N(x'/x_0)\approx 1$ for small values of the argument x'/x_0. In particular, for $x'=\frac{1}{4}$ our chosen form of $\mathfrak{F}_N$ (Appendix A) gives $\mathfrak{F}_N(1/4x_0)\approx 0.96$ for the 2-cm lead case, and 0.93 for the 5-cm lead case. The integral up to κ is thus the same as in the case of the point nucleus.

$$s(\eta)=-\frac{1}{2G}\int_0^\kappa \frac{(\eta x')^2dx'}{2(x'^2+x_m^2)^{\frac{1}{2}}}$$

$$+\frac{1}{2G}\int_\kappa^\infty \frac{\mathfrak{F}_N(x'/x_0)(\cos\eta x'-1)dx'}{x'^3}. \quad (37)$$

Performing the first integration and using the fact that $x_m\ll\kappa$, Eq. (37) becomes

$$s(\eta)=-\frac{\eta^2}{4}-\frac{\eta^2}{4G}\ln\left(\frac{\kappa}{1.26}\right)$$

$$+\frac{1}{2G}\int_\kappa^\infty \frac{\mathfrak{F}_N(x'/x_0)(\cos\eta x'-1)dx'}{x'^3}, \quad (38)$$

where $\ln(1/1.26)=(\ln 2\gamma-\frac{3}{2})$.

It is understood that κ must be much smaller than 1. If now $s(\eta)$ is put into Eq. (36a) and the terms multiplied by $(2G)^{-1}$ are expanded in the same way as before, after the η integrations, the following result is obtained for the first order in $(2G)^{-1}$.

$$M_2(x)=\frac{\exp(-x^2)}{\sqrt{\pi}}\left[1+\frac{1}{4G}2(2x^2-1)\ln(\kappa/1.26)\right]$$

$$+\frac{1}{4G}\frac{1}{\sqrt{\pi}}N(\kappa,x), \quad (39)$$

where

$$N(\kappa,x)=\int_\kappa^\infty \frac{\mathfrak{F}_N(\lambda/x_0)}{\lambda^3}T(x,\lambda)d\lambda,$$

$$T(x,\lambda)=\exp[-(x+\lambda)^2]+\exp[-(x-\lambda)^2]-2\exp(-x^2).$$

In its present form Eq. (39) is convenient for calculation only for large values of $x(x\geq 4$ in the 5 cm lead case) because in this case $\exp(-x^2)$ is so small that the precise value of κ does not influence the result.

However, for smaller values of x the exact value of κ is of importance. In order that the expansion of $(\cos\eta x'-1)$ in Eq. (37) be correct K must be much smaller than 1. However, it is inconvenient to evaluate the final integral of Eq. (39) numerically for very small values of κ.[8]

To overcome this difficulty one can use the property of the form factor that $\mathfrak{F}_N(\lambda/x_0)\approx 1$ for values of λ much larger than η_c^{-1}. Then $\int_\kappa^L\lambda^{-3}T(x,\lambda)d\lambda$ can be evaluated analytically, where $L\approx\frac{1}{4}$. This makes the numerical integration $\int_L^\infty\lambda^{-3}T(x,\lambda)\mathfrak{F}_N(\lambda/x_0)d\lambda$ practicable. In doing this it will be seen that the dependence upon κ vanishes and a convenient and accurate expression for $M_2(x)$ is obtained. We observe that if $L\approx\frac{1}{4}$ then λ can be considered small even if $2x\lambda$ is not. Thus we can expand $\exp(-\lambda^2)$ in $T(x,\lambda)$:

$$T(x,\lambda)=2\exp(-x^2)\exp(-\lambda^2)(\cosh 2\lambda x-1)$$

$$\approx 2\exp(-x^2)(1-\lambda^2+\cdots)(\cosh 2\lambda x-1), \quad (40)$$

if we take only terms up to λ^2. We evaluate

$$J=\int_\kappa^L \lambda^{-3}T(x,\lambda)d\lambda\approx 2\exp(-x^2)$$

$$\times\int_\kappa^L\left[\frac{(\cosh 2\lambda x-1)}{\lambda^3}-\frac{\cosh 2\lambda x}{\lambda}\right]d\lambda. \quad (41)$$

This, after some labor, neglecting terms in κ^2 or higher,

[8] We can see more clearly what error is introduced if κ is allowed to become large by doing the following. Expand $T(x,\lambda)$ of Eq. (39) in powers of $2x\lambda$. Then neglecting terms of order λ^4 or $(x\lambda)^4$ gives $T(x,\lambda)\approx 2\lambda^2\exp(-x^2)(2x^2-1)$. Integrating $N(\kappa,x)$ from κ to L and letting $\mathfrak{F}_N(\lambda/x_0)=1$ in this integral, we obtain

$$M_2(x)=\frac{1}{\sqrt{\pi}}\left[\exp(-x^2)+\frac{1}{4G}2\exp(-x^2)(2x^2-1)\right.$$

$$\left.\times\ln\left(\frac{L}{1.26}\right)+\frac{1}{4G}N(L,x)\right].$$

But this is just the same result obtained by letting $\kappa=L$ in the Fourier transform $s(\eta)$ of Eq. (37). Thus allowing κ to become large introduces the same error as would be introduced if the final result were expanded in powers of $2x\lambda$.

gives

$$J\approx\exp(-x^2)\left[\frac{1}{L^2}-\frac{\cosh 2Lx}{L^2}-\frac{2x}{L}\sinh 2Lx\right]$$
$$+6x^2\exp(-x^2)+2\exp(-x^2)(2x^2-1)$$
$$\times\left[\int_0^{2Lx}\frac{(\cosh t-1)}{t}dt+\ln L\right]$$
$$-2\exp(-x^2)(2x^2-1)\ln\kappa. \quad (42)$$

Putting this result into Eq. (39) gives

$$M_2(x)=\frac{\exp(-x^2)}{\sqrt{\pi}}\left[1+\frac{q(L,x)}{4G}\right]$$
$$+\frac{1}{4G}\frac{1}{\sqrt{\pi}}\int_L^\infty\lambda^{-3}\mathfrak{F}_N(\lambda/x_0)T(x,\lambda)d\lambda, \quad (43)$$

where[9]

$$q(L,x)=2(2x^2-1)\left[\ln\left(\frac{L}{1.26}\right)+\int_0^{2Lx}\frac{(\cosh t-1)}{t}dt\right]$$
$$+6x^2-\frac{1}{L^2}(\cosh 2Lx-1)-\frac{2x}{L}\sinh 2xL,$$

and

$$\int_0^{2Lx}dt\frac{(\cosh t-1)}{t}\approx\frac{(2Lx)^2}{2\cdot 2}+\frac{(2Lx)^4}{4\cdot 4!}+\cdots.$$

D. Evaluation of the Integrals

The integrals $K(x)$ and $N(L,x)$ which occur in expressions for the multiple scattering distributions derived above can easily be evaluated by numerical means. $N(L,x)$ in particular can be evaluated conveniently for values of $L=\frac{1}{4}$ or $\frac{1}{2}$, and for such values of L Eq. (43) is accurate. These integrals do not depend on the momentum of the incoming particle but rather on β, which ≈ 1 in the relativistic region. Thus a single computation is sufficient for all relativistic momenta. For different thicknesses, or atomic number, the form factor is changed only through the change in x_0 in $\mathfrak{F}_N(x/x_0)$. Table I gives the function $T(x,\lambda)\lambda^{-3}$ for various values of x at grid spacings of $\frac{1}{4}$ for λ. Table II gives the values of the function $q(L,x)$ for $L=\frac{1}{4}$ for relevant values of x.

$K(x)$ and $N(L,x)$ have been evaluated numerically for the 2-cm and 5-cm lead cases for grid spacings of $\Delta x=\frac{1}{2}$ and $\frac{1}{4}$, by using Weddle's rule.[10] Comparison of the results for the two grids shows the largest numerical errors occur for the small values of x where the correction term is unimportant. For larger values of x ($x\gtrsim 3$), the change in the numerical results for grid changes from $\frac{1}{2}$ to $\frac{1}{4}$ is less than 2 percent. In all, the errors in the distribution due to numerical errors resulting from the integration, for a grid of $\frac{1}{4}$, appear to be less than 1 percent.

For very large values of x, asymptotic formulas can be developed. For a form factor which decreases asymptotically as const/x^{n-3}, the asymptotic expression is

$$\sigma_M=\sigma_S\left[1+\frac{n(n+1)}{4x^2}+\cdots\right].$$

Such expressions give reasonably accurate results for $x\geq 6$ (5 percent or better). The procedure described near the end of section 2 is also readily adapted to this method for a quick graphical correction procedure to obtain $M_2(x)$ from the single scattering law. For real atoms at the larger angles incoherent scattering predominates so that the form factor becomes $\mathfrak{F}_N(\varphi/\varphi_0)\approx Z^{-1}$. Then the multiple scattering distribution is given by the asymptotic expansion of the Molière distribution multiplied by Z^{-1}. This is

$$M_2(x)dx=\frac{\sigma_s}{Z}\left[1+\frac{3}{x^2}+\frac{45}{4x^4}+\cdots\right].$$

The results presented in this section have included only the first powers in $(2G)^{-1}$. It is possible to obtain the next order, but in view of the large uncertainty in the nuclear form factor $\mathfrak{F}_N(\varphi/\varphi_0)$ it was not considered worthwhile at present to consider this term.

We wish to thank Professor Robert Serber for helpful discussions of the single-proton form factor. We also wish to thank Miss Hilda Oberthal for her assistance with the computations of the results of the modified Molière theory.

APPENDIX A. THE SINGLE-SCATTERING LAW

The single scattering laws used in the calculations were obtained in the following way. The Rutherford formula, modified at small angles due to electron shielding, is given in Eq. (3) following Molière. When nuclear extension is considered, this must be multiplied by a nuclear form factor $\mathfrak{F}_N(\varphi/\varphi_0)$ which contains a part representing elastic coherent scattering plus a part representing inelastic scattering, $\mathfrak{F}_N(\varphi/\varphi_0)=F_N{}^C(\varphi/\varphi_0)+F_N{}^I(\varphi/\varphi_0)$. In principle, if the nuclear wave functions were known exactly, one would prefer to calculate $F_N{}^C$ by an exact phase shift analysis for spin $\frac{1}{2}$ particles, and calculate $F_N{}^I$ by considering in detail all of the possible final states of the scattering system. However, the nuclear charge distribution and the nuclear wave functions are not known exactly, and, in fact, measurements of $F_N{}^C$ for fast electrons[11] are providing valuable

[9] In our earlier communication by the same title, Leon N. Cooper and James Rainwater, Phys. Rev. **95**, 1107 (1954), $q(L,x)$ was given as $2(2x^2-1)q(\kappa)=2(2x^2-1)\ln(\kappa/1.26)$ which was a sufficiently good approximation in the example there considered.

[10] H. Margenau and G. M. Murphy, *The Mathematics of Physics and Chemistry* (D. Van Nostrand Company, Inc., New York, 1943), p. 461.

[11] Pidd, Hammer, and Raka, Phys. Rev. **92**, 436 (1953); Hofstadter, Fechter, and McIntyre, Phys. Rev. **92**, 978 (1953); L. I. Schiff, Phys. Rev. **92**, 988 (1953); Hofstadter, Hahn, Knudsen, and McIntyre, Phys. Rev. **95**, 512 (1954); Yennie, Ravenhall, and Wilson, Phys. Rev. **95**, 500 (1954).

TABLE I. Values of $T(x,\lambda)\lambda^{-3}=\{\exp[-(x+\lambda)^2]+\exp[-(x-\lambda)^2]-2\exp(-x^2)\}\lambda^{-2}$.

λ	$x=0$	$x=1$	$x=2$	$x=3$	$x=4$	$x=5$	$x=6$	$x=7$
			$\times10^{-1}$	$\times10^{-2}$	$\times10^{-5}$	$\times10^{-8}$	$\times10^{-12}$	
0.25	−7.755	2.793	10.540	1.9112	3.650	0.8463	0.2504	
0.50	−3.539	1.1875	5.656	1.3507	3.649	1.2620	0.5793	
0.75	−2.040	0.5936	4.113	1.4421	6.079	3.385	2.537	
1.00	−1.2642	0.2826	3.314	1.8069	12.319	11.251	13.887	
	$\times10^{-1}$	$\times10^{-2}$	$\times10^{-1}$	$\times10^{-2}$	$\times10^{-4}$	$\times10^{-6}$	$\times10^{-9}$	$\times10^{-13}$
1.25	−8.094	10.751	2.730	2.382	2.659	0.3999	0.08138	0.02241
2.50	−5.301	1.3329	2.199	3.116	5.719	1.4178	0.4756	0.2159
1.75	−3.557	−3.087	1.6845	3.907	11.810	4.827	2.670	1.9982
2.00	−2.454	−4.597	1.2042	4.595	22.89	15.426	14.067	17.360
	$\times10^{-2}$	$\times10^{-2}$	$\times10^{-2}$	$\times10^{-2}$	$\times10^{-3}$	$\times10^{-4}$	$\times10^{-7}$	$\times10^{-10}$
2.25	−17.447	−4.619	7.926	5.000	4.106	0.4561	0.6858	0.13953
2.50	−12.775	−4.034	4.7499	4.983	6.746	1.2355	3.062	1.0273
2.75	−9.612	−3.313	2.564	4.516	10.079	3.044	12.438	6.879
3.00	−7.406	−2.657	1.2268	3.703	13.625	6.784	45.71	41.68
	$\times10^{-2}$	$\times10^{-2}$	$\times10^{-3}$	$\times10^{-3}$	$\times10^{-2}$	$\times10^{-3}$	$\times10^{-5}$	$\times10^{-7}$
3.25	−5.826	−2.125	5.039	27.36	1.6598	1.3625	1.5136	17.2276
3.50	−4.665	−1.7116	1.6040	18.159	1.8165	2.458	4.503	1.1161
3.75	−3.793	−1.3942	0.19227	10.800	1.7814	3.975	12.003	4.905
4.00	−3.125	−1.1494	−0.28618	5.744	1.5625	5.748	28.62	19.283
	$\times10^{-2}$	$\times10^{-3}$	$\times10^{-4}$	$\times10^{-4}$	$\times10^{-3}$	$\times10^{-3}$	$\times10^{-4}$	$\times10^{-5}$
4.25	−2.605	−9.584	−3.947	27.27	12.237	7.423	6.093	0.6769
4.50	−2.195	−8.074	−3.808	11.539	8.547	8.547	11.567	2.118
4.75	−1.8662	−6.865	−3.370	4.341	5.317	8.765	19.558	5.906
5.00	−1.6000	−5.886	−2.921	1.4455	2.943	8.000	29.43	14.65
	$\times10^{-3}$	$\times10^{-3}$	$\times10^{-4}$	$\times10^{-6}$	$\times10^{-4}$	$\times10^{-3}$	$\times10^{-3}$	$\times10^{-4}$
5.25	−13.821	−5.085	−2.530	42.04	14.486	6.492	3.924	3.232
5.50	−12.021	−4.422	−2.201	10.120	6.335	4.681	4.681	6.335
5.75	−10.520	−3.870	−1.9268	1.4347	2.460	2.997	4.941	11.026
6.00	−9.259	−3.406	−1.6959	−0.5713	0.8479	1.7031	4.630	17.031
	$\times10^{-3}$	$\times10^{-3}$	$\times10^{-4}$	$\times10^{-7}$	$\times10^{-6}$	$\times10^{-4}$	$\times10^{-3}$	$\times10^{-3}$
6.25	−8.192	−3.014	−1.5004	−9.050	25.93	8.586	3.848	2.326
6.50	−7.283	−2.679	−1.3339	−8.813	7.029	3.838	2.836	2.836
6.75	−6.503	−2.392	−1.1911	−8.000	1.689	1.5208	1.8465	3.055
7.00	−5.831	−2.145	−1.0680	−7.193	0.3591	0.5340	1.0726	2.916
	$\times10^{-3}$	$\times10^{-3}$	$\times10^{-5}$		$\times10^{-8}$	$\times10^{-6}$	$\times10^{-4}$	$\times10^{-3}$
7.25	−5.248	−1.9307	−9.613		6.729	16.610	5.500	2.465
7.50	−4.741	−1.7440	−8.683		1.0809	4.576	2.498	1.8461
7.75	−4.297	−1.5806	−7.869		0.1195	1.1162	1.0048	1.2200
8.00	−3.906	−1.4370	−7.155		⋯	0.24103	0.3577	0.7185
	$\times10^{-3}$	$\times10^{-3}$	$\times10^{-5}$			$\times10^{-9}$	$\times10^{-6}$	$\times10^{-5}$
8.25	−3.562	−1.3103	−6.524			46.07	11.273	37.33
8.50	−3.257	−1.1980	−5.965			7.792	3.143	17.162
8.75	−2.985	−1.0983	−5.468			1.1660	0.7756	6.981
9.00	−2.743	−1.0092	−5.025			0.1543	0.1693	2.512

information concerning the nuclear charge distribution. In particular, mu-mesonic x-ray studies,[12] fast electron scattering,[11] and other experimental results[12,13] which are sensitive to the nuclear charge distribution, all agree that the nuclear charge distribution is more compact than had previously been believed to be the case. For a nuclear charge distribution $\rho(r)=\rho_0$ for $r<R=r_0A^{\frac{1}{3}}$, and $\rho(r)=0$ for $r>R$, the radius R has clear meaning and a best match to the above experiments gives $r_0\approx1.0$ to 1.2×10^{-13} cm for not too small A. Elementary considerations of quantum mechanics show that such a model cannot be strictly correct, and

TABLE II. Values of $q(L,x)$ for $L=\frac{1}{4}$. [Eq. (43).]

x	$q(L,x)$
0	3.238
0.5	1.595
1.0	−3.218
1.5	−11.339
2.0	−22.07
2.5	−34.65
3.0	−47.40
3.5	−57.48
4.0	−60.35

[12] Val. L. Fitch and James Rainwater, Phys. Rev. 92, 789 (1953); L. N. Cooper and E. M. Henley, Phys. Rev. 92, 801 (1953); John A. Wheeler, Phys. Rev. 92, 812 (1953); D. L. Hill and K. W. Ford, Phys. Rev. 94, 1617 and 1630 (1954).

[13] F. Bitter and H. Feshbach, Phys. Rev. 92, 837 (1953); B. G. Jancovici, Phys. Rev. 95, 389 (1954); D. C. Peaslee, Phys. Rev. 95, 717 (1954).

$\rho(r)$ must be a continuous function of r. In this case, although the different experiments are not always sensitive in the same manner to the *shape* of $\rho(r)$, the *equivalent* uniform model r_0 is usually taken to be that value which gives the same $\langle r^2\rangle$ as for the non-uniform $\rho(r)$. Prior to late 1953 the value of r_0 would have been considered to be "well known" and of magnitude 1.4 to 1.5×10^{-13} cm. Thus all of the comparisons[5,6] between the experimental and expected multiple-scattering distributions used this larger value of r_0.

Until recently it has been customary to calculate $F_N{}^C$ by using the Born approximation, which gives a linear superposition of the scattering amplitudes of the individual protons so the scattered intensity is of the form

$$A^2|\langle i|\sum_j \exp(-i\mathbf{q}\cdot\mathbf{r}_j)|i\rangle|^2=Z^2A^2F_N{}^C, \qquad \text{(A1)}$$

where A^2 is the scattering intensity of a single proton, $\mathbf{q}=\mathbf{k}-\mathbf{k}_0$ represents the vector momentum change in scattering, $\mathbf{r}_j$ is the position coordinate of the jth proton in the nucleus, and the evaluation is for the ground state $|i\rangle$ of the nucleus. Thus $(F_N{}^C)^{\frac{1}{2}}$ is just the Fourier transform of the nuclear charge distribution. For a uniform nuclear model, letting $y=\varphi/\varphi_0$, this gives

$$F_N{}^C(y)=\left[\frac{3}{y^3}(\sin y-y\cos y)\right]^2, \qquad \text{(A2)}$$

which gives diffraction minima at $y=4.49$, 7.7, etc. Schiff[11] in particular has considered the form of $F_N{}^C(y)$ in Born approximation for various simple analytic forms for $\rho(r)$. We note that $F_N{}^C$ is of the form of the square or a real amplitude term, and thus gives diffraction minima when the amplitude changes sign. For distributions sufficiently peaked in the center and with a gradual "tailing-off," the amplitude may not change sign and $F_N{}^C$ will be a smoothly decreasing function of y. It is readily seen from Eq. (A1) that $F_N{}^C(0)=1$ and, by expanding the exponentials and using inversion symmetry through $r=0$, that the leading term in the decrease of $F_N{}^C$ for small y depends on $\langle r^2\rangle$. In this connection it is interesting to note that the widely used Williams[1] formula uses $V(r)=(Ze^2/r)(1-e^{-2r/R})$ for which $\rho(r)=(4\rho_0/3)(R/r)e^{-2r/R}$, and $\langle r^2\rangle=3R^2/2$. Here ρ_0 is the charge density for a uniform nucleus of radius R. Although Williams implied that this distribution closely approximates a uniform model, and it has been taken by others[6] as corresponding to the uniform model, it actually corresponds to a rather strongly peaked distribution about $r=0$ with $\langle r^2\rangle$ larger by a factor of 5/2 than for a uniform distribution, and thus a correspondingly more rapid initial decrease of $F_N{}^C$ for small y. For larger y it gives a uniform decrease of $F_N{}^C$ with y with an asymptotic form $16/y^4$ for large y. In the region of large y Eq. (A2), between minima, has a steady decrease with y which can be estimated by neglecting the $\sin y$ term and setting $|\cos y|\approx1$ to give an asymptotic form $9/y^4$, which is below that for the Williams distribution.

The detailed phase shift calculations of Yennie, Ravenhall, and Wilson, using various assumed $\rho(r)$, show that the shape of $F_N{}^C(y)$ is energy-dependent and significantly different from the Born approximation value for high-Z materials. In particular, the scattering amplitude is a complex number which circles the value zero in the complex plane when "changing sign." Thus the "diffraction minima" are largely missing, or are greatly reduced in magnitude compared to the Born approximation results. This feature is also apparent in the experimental results and had led to an initial "Born approximation interpretation" that $\rho(r)$ must resemble an exponential distribution. The interpretation favored at the time of this writing is that $\rho(r)$ can be fairly constant for values of r containing most of the charge distribution, with a gradual dropping off at the "surface."

In view of the above results, and with the consideration that we wish to choose a form for $F_N{}^C$ which will *not underestimate* the expected multiple scattering due to the known electromagnetic interaction of mu mesons with protons in examining the results of experiments investigating possible "anomalous scattering," we have chosen the following form for $F_N{}^C$. For $y=0$, 1, 2, 3 we choose $F_N{}^C=1.00$, 0.82, 0.50, and 0.15 to approximate Eq. (A2) after the effect of the first diffraction minimum is "removed." A smooth curve through these points is then joined smoothly to $F_N{}^C=12/y^4$ for $y\gtrsim4$, this being between the Williams formula and the value obtained above for a uniform distribution. In principle this should be applied to the cross section for total angle scattering rather than for projected angle scattering as we do here. The consequences of this approximation are discussed following the discussion of $F_N{}^I$.

The calculation of the inelastic scattering is quite difficult to perform exactly. Amaldi, Fidecaro, and Mariani[14] have carried out detailed calculations of the inelastic and elastic scattering expected, using a particular independent particle nuclear model. For incident particles of very high kinetic energy and $cp\sim100$ Mev, a Born approximation treatment similar to that used in the theory of x-ray scattering by atoms should be reasonably accurate and gives the result, as for x-rays,[15] that $F_N{}^I\approx Z^{-1}(1-F_N{}^C)$ times the form factor for the scattering by a single proton. This last factor must be included if the proton is not effectively a point charge (due, say, to meson-cloud effects). Amaldi *et al.*[14] have particularly emphasized this point and have calculated the expected effect of the proton "size" on the basis of a simple model. Experimentally, however, it seems that the proton charge distribution should be treated as being confined to a surprisingly small volume on the

[14] Amaldi, Fidecaro, and Mariani, Nuovo cimento **7**, 553 (1950).
[15] See A. H. Compton and S. K. Allison, *X-rays in Theory and Experiment* (D. Van Nostrand and Company, Inc., New York, 1935), Chap. 3.

basis of electron scattering experiments,[16] and, by inference, from the interpretation[17] of the experiments on the neutron-electron interaction. The experiments show that the proton form factor is essentially unity for $cp\sim 200$ Mev and at angles $\sim 90°$, so we take it to be essentially unity for $cp\varphi$ values considered in this paper. For large enough values of $cp\varphi$ this factor will eventually become important and require consideration.

If the single-proton form factor is set equal to unity, the above expression for $F_N{}^I$ can be understood by analogy with x-ray scattering where the "inelastic" scattering corresponds to modified Compton elastic scattering, where the recoil momentum is taken up by a single (moving) electron rather than by the atom as a whole. The final states of importance correspond to recoil electron momenta centered about the photon momentum transfer, modified by the initial electron momentum distribution, which is given by the Fourier transform of the ground state wave function of the atom. For charged particles of kinetic energy and $cp \gg 100$ Mev, and for $y \gg 1$, the final states of importance should be attainable with moderate energy loss, so the expression for the total scattering intensity in a given direction can be written with fair accuracy by summing over-all final states, keeping $\mathbf{q}$, for a given angle, the same as in the elastic case.

$$Z^2[F_N{}^C+F_N{}^I] = A^2\sum_f |\langle f|\sum_j \exp(-i\mathbf{q}\cdot\mathbf{r}_j)|i\rangle|^2 = A^2\langle i|\sum_m \exp(i\mathbf{q}\cdot\mathbf{r}_m)\sum_j \exp(-i\mathbf{q}\cdot\mathbf{r}_j)|i\rangle \quad \text{(A3)}$$

by closure. For an independent particle model Eq. (A1) and Eq. (A3) lead to the above quoted result for $F_N{}^I$. Thus we set

$$\mathfrak{F}_N = F_N{}^C+F_N{}^I = F_N{}^C+Z^{-1}(1-F_N{}^C). \quad \text{(A4)}$$

The remaining point to be considered is the error introduced when $\mathfrak{F}_N$ is applied to the law for projected (rather than total) angle scattering. For a $d\sigma/d\Omega$ law varying as φ^{-n}, the projected angle single scattering law is obtained by multiplying $d\sigma/d\Omega$ by $\varphi c(n)$, where $c(n)=\pi/2$, 4/3, 16/15, $5\pi/16$, and 32/35 for $n=4$, 5, 7, 8, and 9 respectively $[c(n)=2\int_0^{\pi/2}\cos^{n-2}\varphi d\varphi]$. In going from $n=4$ for Rutherford scattering to $n=8$, which is obtained when the asymptotic form of $F_N{}^C$ multiplies the Rutherford scattering, $c(n)$ is reduced by the factor of 5/8. This represents an extreme situation since $\mathfrak{F}_N$ is more slowly varying. We note that the above effect can approximately be taken into account by choosing $r_0\sim 10$ percent larger than otherwise when applying $\mathfrak{F}_N$ to the law for projected angle scattering. In the examples we choose $r_0=1.1\times 10^{-13}$ cm, corresponding to $r_0\approx 1.0\times 10^{-13}$ cm for $d\sigma/d\Omega$. In view of the uncertainty concerning the exact form for the true $\mathfrak{F}_N(y)$, we consider this approximation to be adequate for the present.

APPENDIX B. DISCUSSION OF THE EXPERIMENTS[5,6] ON MU-MESON SCATTERING

The interest of the authors in the multiple-scattering theory was mainly stimulated by the possible consequences of a strong anomalous mu-meson–nucleon interaction on the interpretation of the mu-mesonic x-ray experiments.[12] Cosmic-ray experiments on the scattering of fast mu mesons by nuclei suggested that some anomalous scattering exists, but the interpretation of the experiments are not completely unambiguous, and there is considerable disagreement on the magnitude and existence of the effect. When these experiments were analyzed, the expected multiple-scattering distribution for an extended nucleus was obtained by various approximations such as comparing the experimental results with the predictions of the Molière and Olbert theories, or by using the Williams theory. In all cases the old "large" nuclear size was used which gives $\langle r^2\rangle$ twice that favored by recent experiments. Thus the experiments were always analyzed on a basis that underestimated the Coulomb multiple scattering. Aside from the results of the experiments discussed below, we note that the mu-mesonic x-ray results indicate that any anomalous energy independent nuclear potential for the mu meson can in its effect at most be equivalent to a slight change in the choice of the nuclear radius when calculating the Coulomb interaction. This could not explain any significant portion of the anomalous scattering reported in some of the experiments. Also, experiments on the nuclear interaction by Annis *et al.*[18] can be explained without invoking any anomalous interaction. Such an "anomalous interaction" would have to be strongly energy-dependent and thus only appear strongly at high energies to explain the low-energy experiments.

Amaldi and Fidecaro[5] investigated the large-angle scattering of fast mu mesons in the energy bands 200 Mev to 320 Mev and >320 Mev, using a counter hodoscope. They compared the large-angle multiple scattering in iron and lead, emphasizing the iron results as far as anomalous scattering is concerned. About 5×10^5 incident mesons were counted and results for iron were consistent with no anomalous interaction. In their series 2 run on iron 249, 168 particles were incident and 3 scattered particles were observed in each energy band. When an extra 200 g/cm² of bricks were placed above the apparatus to decrease the number of protons etc., 204, 914 particles were incident with one scattered count in the lower energy band and none in the higher energy band. The latter numbers are about the expected values for scattered protons, etc., with that amount of filtering. They set an upper limit for "anomalous scattering" of $\sim 4.5\times 10^{-29}$ cm²/nucleon in the lower energy band and $\sim 2.3\times 10^{-29}$ cm²/nucleon in the upper energy

[16] J. A. McIntyre and R. Hofstader, Bull. Am. Phys. Soc. **29**, No. 6, 19 (1954); Hofstadter, McAllister, and Wiener, Bull. Am. Phys. Soc. **29**, No. 6, 19 (1954).

[17] See L. I. Foldy, Phys. Rev. **87**, 675 (1952) for a discussion of this subject and for reference to earlier papers.

[18] Annis, Wilkins, and Miller, Phys. Rev. **94**, 1038 (1954).

band. These values assume isotropic scattering for the anomalous part.

Whittemore and Shutt investigated the multiple scattering of negative mu mesons in 5 cm of lead for particles having 0.3 Bev $\leqslant cp \leqslant$ 3.1 Bev using two cloud chambers and a magnet for momentum analysis. Their experimental points essentially fall on the Scott-Snyder (Molière) curve for $cp\varphi \lesssim 13$ Bev degrees, with one point at 17 Bev degrees a factor of about two below the Molière curve (but with a large statistical uncertainty). From Fig. 5 these points would also fall above our theoretical curve. The principle difficulty in the interpretation of these results is the question of the uncertainty that only mu mesons were involved. It is of interest in this connection that the later experiments were conducted underground, or, in one case, using 1 meter of lead absorber above the cloud chamber to assure greater beam purity. Also we should like to point out that measurements[19] on the scattering of $cp \sim 200$ Mev π mesons on Pb show differential cross sections for elastic plus inelastic scattering which are always $\sim$0.1 barn or larger. By contrast, the elastic scattering experiments using electrons of comparable momentum give differential cross sections which are $\sim 10^{-6}$ barn at 120°, a factor of 10^5 below the π-meson cross section. This contrast is admittedly extreme, but it emphasizes that one should not underestimate the the possible importance of small percentages of beam contamination in such experiments.

The remaining experiments used greater absorber thickness at sea level, or operated underground to minimize beam contamination. They have the common feature, however, that the momentum distribution of the incident particles was not measured directly, but was assumed known from other sources. Any error in the assumed known momentum distribution would tend to affect all of these experiments in a similar fashion. The fact that the observed scattering distributions are not given directly as a function of $p\varphi$ makes comparison of their final curves with our calculated $M(\varphi)$ difficult.

George, Redding, and Trent measured the multiple scattering of penetrating cosmic-ray particles in 2-cm lead plates at 60 m.w.e. (meters water equivalent) underground using a counter-triggered cloud chamber. Three experimental arrangements were used for the triggering counter telescope involving 0, 5, and 10 cm lead below the cloud chamber. The particles were all assumed to be μ mesons and the momenta of the individual particles were known only to be above the cutoff values determined by the lead absorber thickness. The analysis was made by assuming that the energy distribution was flat for $E \ll E_0 = 12$ Bev. The experimental distribution $N(\varphi)$ was compared with one calculated using a weighted average of Gaussian functions (one for each energy) in accord with the above prescription for the assumed energy distribution. No anomalous scattering was observed using 10 cm of lead, but a small amount of "anomalous scattering" appeared at larger angles when 0 or 5 cm lead was used. They conclude that no anomalous scattering is observed for kinetic energies $> \sim 200$ Mev, and some may exist for lower energies. It is interesting to note that the curves for 0 and 5 cm of lead show essentially the same excess over their theoretical curves whether point-nucleus scattering is included or not. It would be desirable to have a direct measurement of the momentum distribution of the incident particles.

Leontic and Wolfendale used a multiplate cloud chamber at sea level with a counter telescope that required that detected particles traverse 1 meter of lead above the chamber, six 2-cm lead plates inside the chamber, and 0, 5, or 10 cm of lead below the chamber. It was assumed that the 1 meter of lead excluded all but μ mesons from the measurements. The main analysis compared the maximum scattering in any of the center four plates with the rms angle for the four plates. The analysis made the assumption that the basic multiple-scattering law for any given particle should effectively be a Gaussian at all angles if no anomalous scattering were present. By an ingenious analysis they showed that the results were inconsistent with this assumption. A further analysis assumed that an "anomalous scattering Gaussian" would be superimposed on the normal multiple scattering Gaussian in some small fraction of the plate traversals. They then obtained a best matching of parameters. Unfortunately, the selection criterion for considered events was biased in favor of selecting cases where "anomalous scattering" (or the non-Gaussian multiple-scattering tail) occurred. Dr. Wolfendale informs[20] us that a preliminary reexamination suggests that the essential features of the results will not be seriously altered on correcting this bias factor. We have not seen the details of this reexamination.

The measurements of Leontic and Wolfendale have been extended by McDiarmid[20] using a different and interesting analysis. The results seem to contradict those of George, Redding, and Trent in that no anomalous scattering is observed for low energies. At higher energies the experimental results are between the Molière and Olbert multiple scattering curves for a (partially) assumed distribution of incident particle energies. The results are closer to the Molière than to the Olbert curves. We have not, however, performed the detailed folding together of our final curve with their derived momentum distributions as is necessary for a quantitative comparison.

References to earlier papers will be found in the articles discussed above.

[19] John O. Kessler and Leon M. Lederman, Phys. Rev. **94**, 689 (1954).

[20] I. B. McDiarmid, (private communication). We wish to thank Professor G. D. Rochester and Dr. A. W. Wolfendale for corresponding with us concerning the analysis of the experiments of the Manchester group on μ-meson scattering. The paper by McDiarmid is scheduled for publication in the *Philosophical Magazine*.

Gauge Invariance and Renormalization in Electrodynamics

DOMINIQUE RIVIER
Lausanne University, Lausanne, Switzerland, and Division of Physics, National Research Council, Ottawa, Canada
(Received September 13, 1954)

A one-parameter infinitesimal group of gauge transformations for the electromagnetic potential has been defined. On defining the electromagnetic potential by a limiting process from a vector field described by two real components, it is possible to set up a Lagrangian which is invariant under infinitesimal gauge and Lorentz transformations and leads to equations of motions which now include the "supplementary conditions." The constants of motion are gauge-dependent and changes in these constants upon gauge transformation may be regarded as a "renormalization."

THE problem of gauge invariance in electrodynamics has not yet been given a fully satisfactory solution.[1] Although a fundamental inconsistency is removed by using a modified Hilbert space with an indefinite metric,[2] the theory is still in a form which is not gauge-invariant.[3] There are mainly two reasons for such an unsatisfactory situation, namely the existence of the so-called supplementary conditions,[4] and the fact that only finite gauge transformations are well defined for the electromagnetic potential A_μ. More precisely, there is so far no way of defining an infinitesimal group of gauge transformations for A_μ. This is unsatisfactory inasmuch as the existence of such a group is strongly suggested by the experimental fact of charge conservation; for it may be shown that in the case of a free matter field, charge conservation is, under certain conditions, a consequence of the invariance of the equations of motion under infinitesimal gauge transformations.[5] A new way of dealing with the two difficulties mentioned above is sketched below.

First, it is possible to define a one-parameter infinitesimal group of gauge transformations for the electromagnetic potential A_μ in the following way. Besides the gauge variable C independent of space-time coordinates x^μ, one introduces a gauge function $g(C,x^\mu)$, dependent on gauge and space-time variables. A finite gauge transformation of A_μ is given by

$$A_\mu \rightarrow A_\mu{}^* = A_\mu + d_\mu g, \tag{1}$$

and an infinitesimal transformation by

$$A_\mu \rightarrow A_\mu{}^* = (1-i\delta C q)A_\mu = A_\mu + \delta C d_\mu g_C', \qquad g_C' \equiv dg/dC. \tag{2}$$

$q \equiv id/dC$ is the infinitesimal operator of the gauge group. To keep the theory in the framework of the special theory of relativity, one assumes that

$$[q,x^\mu]_-=0, \quad [q,p_\mu]_-=0, \quad [q,m_{\mu\nu}]_-=0, \tag{3}$$

where p_μ and $m_{\mu\nu}$ are the infinitesimal operators of the Lorentz group. The matter field described by the spinor ψ has the following gauge dependence:

$$q\psi = \epsilon g_C'\psi, \quad q\bar{\psi} = -\epsilon g_C'\bar{\psi}, \quad \epsilon \equiv e/\hbar c. \tag{4}$$

The theory is said to be gauge-invariant if the equations of motion are in a form which is independent of the special choice for the gauge, i.e., invariant under an infinitesimal gauge transformation defined by (2) and (4).

Secondly, concerning the presence of supplementary condition, it is known that one way of avoiding it is to consider A_μ as a neutral field corresponding to quanta having a very small rest mass $\kappa \neq 0$.[6] Unfortunately, if $\kappa \neq 0$ there is no gauge dependence of the type (3) for A_μ (in fact, A_μ is assumed to be gauge independent in these theories), and this fact seems to show an essential difference between the case $\kappa=0$ and the case $\kappa \neq 0$.[7] The situation is different, however, if one considers a vector field described by *two real* components $A_\mu{}^{(1)}$ and $A_\mu{}^{(2)}$, depending on the gauge according to

$$qA_\mu{}^{(1)} = i\epsilon(\kappa/K)A_\mu{}^{(2)}, \quad qA_\mu{}^{(2)} = -i\epsilon(\kappa/K)A_\mu{}^{(1)}, \tag{5}$$

where K is the mass of the electron and κ the mass of the quantum associated with the fields $A_\mu{}^{(\alpha)}$. One may define then the electromagnetic potential by the limiting process:

$$\lim_{\kappa\to 0} A_\mu{}^{(1)} = A_\mu, \quad \lim_{\kappa\to 0} \epsilon(\kappa/K)A_\mu{}^{(2)} = d_\mu g_C'. \tag{6}$$

By using this method, it is possible to set up a Lagrangian of the form

$$L = \sum_{\alpha=1}^{2}\left(\tfrac{1}{4}F^{\mu\nu(\alpha)}F_{\mu\nu}{}^{(\alpha)} - \frac{\kappa^2}{2}A^{\lambda(\alpha)}A_\lambda{}^{(\alpha)}\right) + \hbar c K^3\bar{\psi}\{\tfrac{1}{2}(id_\mu{}^{LA}\gamma^\mu - \gamma^\mu id_\mu{}^{A}) + K\}\psi, \tag{7}$$

$$id_\mu{}^{A} \equiv id_\mu + \sum_{\alpha=1}^{2}\xi^{(\alpha)}A_\mu{}^{(\alpha)}, \quad id_\mu{}^{LA} \equiv id_\mu{}^{L} + \sum_{\alpha=1}^{2}\xi^{L(\alpha)}A_\mu{}^{(\alpha)}$$

[1] S. T. Ma, Phys. Rev. **80**, 729 (1950); F. Coester and J. M. Jauch, Phys. Rev. **78**, 149 (1950); H. Snyder, Phys. Rev. **87**, 164 (1952).
[2] S. N. Gupta, Proc. Phys. Soc. (London) **A63**, 681 (1950); K. Bleuler, Helv. Phys. Acta **23**, 567 (1950).
[3] G. Källén, Helv. Phys. Acta **25**, 417 (1950), p. 420.
[4] For instance, F. Coester and J. M. Jauch, reference 1.
[5] J. Schwinger, Phys. Rev. **82**, 914 (1951); D. Rivier, Progr. Theoret. Phys. (Japan) **9**, 633 (1953); the notation of the latter paper is used here.
[6] L. de Broglie, *Une nouvelle théorie de la lumière* (Hermann, Paris, 1940–1942); E. C. G. Stueckelberg and T. A. Green, Helv. Phys. Acta **24**, 153 (1951).
[7] W. Pauli, Revs. Modern Phys. **13**, 203 (1941), p. 219.

which is invariant under an infinitesimal gauge transformation and a homogeneous Lorentz transformation and which leads, for the limit $\alpha\to 0$, to the equations of motions

$$-d_\lambda F^{\lambda\mu}+J_0{}^\mu=0,\quad d_\lambda A^\lambda=0,\quad \Box g_C'=0,$$
$$\bar\psi\{(id^{\lambda L}+\epsilon A^\lambda)\gamma_\lambda+K\}=0, \tag{8}$$
$$\{(-id^\lambda+\epsilon A^\lambda)\gamma_\lambda+K\}\psi=0,$$

with

$$F^{\lambda\mu}\equiv d^\lambda A^\mu-d^\mu A^\lambda,\quad J_0{}^\mu\equiv eK^3\bar\psi\gamma^\mu\psi,$$

i.e., the equations of motions of electrodynamics together with the supplementary conditions

$$d_\lambda A^\lambda=0,\quad \Box g_C'=0,\quad \text{whence}\quad \Box g=0, \tag{9}$$

which appear now as equations of motion. The gauge function $g(C,x)$ is treated on the same footing as the fields A_μ and ψ.

The constants of motion are obtained by using Cartan's method.[5] They are the total charge:

$$Q=\int_\Sigma d\sigma_\lambda J^\lambda,\quad J^\lambda\equiv J_0{}^\lambda g_C'-F^{\lambda\rho}d_\rho g_C'\,; \tag{10}$$

the total energy momentum vector:

$$P_\mu=\int_\Sigma d\sigma_\lambda T_\mu{}^\lambda,$$
$$T_\mu{}^\lambda\equiv g_\mu{}^\lambda L_0-F^{\lambda\rho}d_\mu A_\rho+(K/\epsilon)^2 d^\lambda g_C' d_\mu g_C' \tag{11}$$
$$-\hbar cK^3\bar\psi\tfrac{1}{2}(id_\mu{}^L\gamma^\lambda-\gamma^\lambda id_\mu)\psi,$$

with

$$L_0=\lim_{\kappa\to 0}L;$$

the total angular momentum tensor:

$$M_{\mu\nu}=\int_\Sigma d\sigma_\lambda Z^\lambda{}_{\mu\nu},$$
$$Z^\lambda{}_{\mu\nu}\equiv x_\mu T^\lambda{}_\nu-x_\nu T^\lambda{}_\mu-(F^\lambda{}_\mu A_\nu-F^\lambda{}_\nu A_\mu) \tag{12}$$
$$+\hbar cK^3\tfrac{1}{4}i\bar\psi(\gamma_\mu\gamma_\nu\gamma^\lambda-\gamma^\lambda\gamma_\mu\gamma_\nu)\psi.$$

Characteristic of the theory is the fact that the constants of motion are gauge-dependent: one has

$$\frac{d}{dC}Q=-\int_\Sigma d\sigma_\lambda d_\rho(F^{\lambda\rho}g_{CC'}),$$
$$\frac{d}{dC}P_\mu=-\int_\Sigma d\sigma_\lambda d_\rho(F^{\lambda\rho}d_\mu g_C'), \tag{13}$$
$$\frac{d}{dC}M_{\mu\nu}=-\int_\Sigma d\sigma_\lambda d_\rho[F^{\lambda\rho}(x_\mu d_\nu-x_\nu d_\mu)g_C'].$$

It is important to notice that each quantity dQ/dC, dP_μ/dC, $dM_{\mu\nu}/dC$ is itself a constant of motion.

One sees therefore that to a gauge transformation corresponds a well defined change in the values of the constants of motion (and consequently of the total mass

$$M=+(P_\mu P^\mu)^{\frac{1}{2}} \tag{14}$$

of the system). This change may be considered as an unobservable renormalization of the constants of motion. The theory itself is gauge-invariant in the sense given above. Details of this work will be published elsewhere.

I would like to express my gratitude to Dr. G. Herzberg for his kind hospitality and to the National Research Council for research facilities. It is a pleasure also to thank Dr. T. Y. Wu and members of the theoretical physics group for discussions.

Quantum-Mechanical Operators

V. ROJANSKY
Union College, Schenectady, New York

(Received May 18, 1954; revised manuscript received October 5, 1954)

The eigenvalue spectra of the Schrödinger coordinate operators, such as x, are continuous. Hellund and Tanaka have recently proposed to replace these operators by operators having discrete eigenvalues. In this note the question is raised whether the operators associated with the coordinates of a particle should have any eigenvalues at all.

HELLUND and Tanaka,[1] using as a background the work of Snyder on quantized space-time,[2] have recently proposed to replace some of the Schrödinger operators by certain other operators. It may therefore be useful to mention at this time another approach to quantum-mechanical operators, even though it is not yet clear to the writer whether this approach is sound.

A precise measurement of the position of a particle cannot be made with a microscope.[3] We hazard the guess that such a measurement cannot be made with any apparatus. We then recall that the eigenvalues of the operator associated with a dynamical variable are the possible results of a precise measurement of this variable. And we are thus led to conclude that the operator X, which is to be associated with what we may vaguely call the x coordinate of a particle, must have no eigenvalues. This argument, extended to include y, z, and perhaps also t, is the main point of this note.

To indicate a method of search for X, Y, Z, and T, we consider the artificial case of X alone and look for an operator that resembles the Schrödinger x but has no eigenvalues. Take a bell-shaped function $C(s-x)$ of s that has a maximum at $s=x$, integrates to unity in the infinite interval, and is different from the δ function; define the operator U by the equation

$$Uf(x)=\int_{-\infty}^{\infty} C(s-x)f(s)ds;$$

and let

$$X=\tfrac{1}{2}(xU+Ux).$$

The function e^{iax} is an eigenfunction of U for every real a and hence U depends only on the Schrödinger operator p_x, which we keep as the momentum operator but for uniformity denote by P_x. Consequently P_x commutes with U and therefore $XP_x-P_xX=ihU$. The eigenvalues of U range from 1 (for zero momentum) to 0 (for infinite momentum). The question of the existence of eigenvalues of X involves a study of $C(s-x)$. In contrast to the work of Foldy and Wouthuysen,[4] we deny the possibility of a precise measurement of the position of a particle, and hence neither our x nor our X can be properly regarded as a "mean-position operator" of a particle.

If the origin is translated in such a way that $x\rightarrow x+x_0$, we have $X\rightarrow X+x_0U$, so that, unless the particle is at rest, X undergoes a contracted displacement. The contraction factor can be made equal to the Lorentz factor; in fact,

$$U=[1+(P_x/mc)^2]^{-\frac{1}{2}}$$

if

$$C(s-x)=(\pi\lambda)^{-1}K_0(\lambda^{-1}|s-x|),$$

where $\lambda=\hbar/mc$ and K_0 is the Bessel function of the third kind and order zero.[5]

The writer has not been able to construct a generalized kernel for U that leads to plausible operators X, Y, Z, and T. If we restrict ourselves to spatial coordinates and small speeds, and proceed in what appears to be the simplest way, we get the approximate formula $U=1+\frac{1}{2}l^2\nabla^2$, so that $X=x+\frac{1}{2}l^2(x\nabla^2+\partial/\partial x)$, with similar results for Y and Z; here l is an unspecified length, probably of the order of magnitude of the Compton wavelength of the particle in question. Approximate coordinate commutators, such as $XY-YX=i(l^2/\hbar)m_z$, agree with those given by Hellund and Tanaka, but the operators X, Y, and Z themselves do not. The use of coordinate operators involving derivatives implies of course a reinterpretation of $\bar{\psi}\psi$.

I am indebted to Messrs. Gregory Breit, A. T. Goble, and Frank Stern for discussions.

[1] E. J. Hellund and K. Tanaka, Phys. Rev. **94**, 192 (1954).
[2] H. S. Snyder, Phys. Rev. **71**, 38 (1947); **72**, 68 (1947).
[3] W. Pauli, *Handbuch der Physik* (Springer, Berlin, 1933), Vol. 24, Part 1, p. 91.
[4] L. L. Foldy and S. A. Wouthuysen, Phys. Rev. **78**, 29 (1950).
[5] G. N. Watson, *Theory of Bessel Functions* (Cambridge University Press, London, 1922 and 1944), p. 388, Eq. (10). G. A. Campbell and R. M. Foster, *Fourier Integrals for Practical Applications* (Bell Telephone System, Monograph B-584, New York, 1931), p. 60, Eq. (558).

Use of Causality Conditions in Quantum Theory

MARVIN L. GOLDBERGER
Institute for Nuclear Studies and Department of Physics, University of Chicago, Chicago, Illinois
(Received October 8, 1954)

It is shown that the derivation of the Kramers-Kronig dispersion relations given by Gell-Mann, Goldberger, and Thirring may be carried out without the use of perturbation theory. It is further shown that the results are essentially independent of the form of the coupling between the electromagnetic field and the matter fields. The demonstration is facilitated by the construction of simple rigorous expressions for the commutator of two vector potential operators and for the S-matrix describing photon scattering.

I. INTRODUCTION

IN a recent paper[1] with the above title, it was shown that the well-known Kramers-Kronig dispersion relations followed from conventional quantum electrodynamics by the imposition of the requirement that signals cannot propagate faster than the velocity of light. The situation discussed was that of a system interacting with the electromagnetic field according to the interaction Hamiltonian, H, given by

$$H=-\int d^3x j_\mu(x)A_\mu(x). \tag{1.1}$$

The proof of the dispersion relations was carried out by treating the above interaction as a perturbation and only terms of second order in the electric charge were retained. It was conjectured that this perturbation limitation was unnecessary and that the results were rigorously correct. It is one of the purposes of this paper to prove that conjecture. We shall also show that the electromagnetic interaction with the matter system need not be restricted to the above simple form which is not even sufficiently general to include the practical case of a system of nucleons and pions. We need, in fact, assume only that the current density operator, defined symbolically by

$$j_\mu(x)=\partial L/\partial A_\mu(x), \tag{1.2}$$

where L is the Lagrangian density of the system and $A_\mu(x)$ is the vector potential operator, does not contain any time derivatives of $A_\mu(x)$. The explicit form of the interaction Hamiltonian will not be needed.

In Sec. II an exact evaluation of the commutator of two A's will be given and in Sec. III we present a rigorous expression for the photon scattering amplitude. The results are briefly discussed in Sec. IV.

II. EVALUATION OF $[A_\mu(x), A_\nu(y)]$

The causality condition used in GGT was the requirement that the commutator of two Heisenberg field operators of the electromagnetic field, $A_\mu(x)$ and $A_\nu(y)$, shall be zero if the points x and y have space-like separation. It was in the evaluation of this commutator that perturbation theory was employed in GGT and only terms of order e^2 were retained in the case when the interaction was given by Eq. (1.1). We shall show that the expression derived there, GGT, Eqs. (3.1) and (3.6), are rigorously correct provided the interaction representation current density operators in those equations are replaced by the corresponding Heisenberg representation operators. There are, of course, other modifications, to be given below, when electromagnetic interactions more general than (1.1) are permitted.

The unrenormalized vector potential operator is taken to satisfy the Heisenberg equation of motion,

$$\Box^2 A_\mu(x)=-j_\mu(x), \tag{2.1}$$

where we assume that the matter system current density operator, $j_\mu(x)$, may involve $A_\mu(x)$ (as is the case for spin-zero fields) but does not contain any time derivations of $A_\mu(x)$. For our purposes there is no need to use the explicitly renormalized operators.[2] We write, following Yang and Feldman[3] and Källén,[4] in place of (2.1) the formal solution

$$A_\mu(x)=A_\mu^{\text{in}}(x)+\int d^4x' D_r(x-x')j_\mu(x'), \tag{2.2}$$

where $D_r(x)$ is the usual retarded Green's function. The "in" fields, $A_\mu^{\text{in}}(x)$, satisfy the commutation relation

$$[A_\mu^{\text{in}}(x),A_\nu^{\text{in}}(y)]=i\delta_{\mu\nu}D(x-y). \tag{2.3}$$

If $\eta(x)$ is the step function which is zero for $x_0<0$ and unity for $x_0>0$, we have

$$D_r(x)=-\eta(x)D(x). \tag{2.4}$$

We now form the commutator of two A's, using

[1] Gell-Mann, Goldberger, and Thirring, Phys. Rev. **95**, 1612 (1954). This paper will be referred to hereafter as GGT.

[2] It is of course possible to use the renormalized operators, but the subsequent equations become more lengthy, since $j_\mu(x)$ then involves time derivatives of $A_\mu(x)$. See, for example, S. N. Gupta, Proc. Phys. Soc. (London) **A64**, 426 (1951). We shall note the modifications of our formulae which would arise in this case.

[3] C. N. Yang and D. Feldman, Phys. Rev. **79**, 972 (1950).

[4] G. Källén, Arkiv Fysik **2**, 371 (1950).

Eq. (2.2):

$$[A_\mu(x),A_\nu(y)]$$
$$=i\delta_{\mu\nu}D(x-y)+\int d^4x'\int d^4y' D_\tau(x-x')$$
$$\times[j_\mu(x'),j_\nu(y')]D_\tau(y-y')$$
$$+\int d^4y' D_\tau(y-y')[A_\mu^{\rm in}(x),j_\nu(y')]$$
$$+\int d^4x' D_\tau(x-x')[j_\mu(x'),A_\nu^{\rm in}(y)]. \quad (2.5)$$

In order to evaluate the commutator of the j's with the "in" fields, we use the following convenient expression given by Källén[5] for $A_\mu^{\rm in}(x)$:

$$A_\mu^{\rm in}(x)=\int d^4x'\Big\{D(x-x')\eta(z-x')j_\mu(x')-\delta(x_0'-z_0)$$
$$\times\Big[A_\mu(x')\frac{\partial D(x-x')}{\partial x_0}+\frac{\partial A_\mu(x')}{\partial x_0'}D(x-x')\Big]\Big\}. \quad (2.6)$$

In this expression z is an arbitrary point. In order to evaluate $[A_\mu^{\rm in}(x),j_\nu(y')]$ we choose z to coincide with y' and similarly in the evaluation of $[j_\mu(x'),A_\nu^{\rm in}(y)]$, z is taken to be x'. We find then for Eq. (2.5) the expression

$$[A_\mu(x),A_\nu(y)]$$
$$=i\delta_{\mu\nu}D(x-y)+\int d^4x'\int d^4y'\Big\{D_\tau(x-x')D_\tau(y-y')$$
$$\times[j_\mu(x'),j_\nu(y')]+D_\tau(y-y')D(x-x')$$
$$\times\Big[\eta(y'-x')[j_\mu(x'),j_\nu(y')]-\delta(x_0'-y_0')$$
$$\times\Big[\frac{\partial A_\mu(x')}{\partial x_0'},j_\nu(y')\Big]\Big]+D_\tau(x-x')D(y-y')$$
$$\times\Big[\eta(x'-y')[j_\mu(x'),j_\nu(y')]$$
$$-\delta(x_0'-y_0')\Big[j_\mu(x'),\frac{\partial A_\nu(y')}{\partial y_0'}\Big]\Big]\Big\}. \quad (2.7)$$

With the exception of the terms involving the commutator of the current density and the time derivative of the vector potential, this is exactly the expression found in GGT, Eq. (3.6), provided that one replaces their interaction representation operators by our Heisenberg operators, $j_\mu(x)$. The additional commutators in (2.7) are of course zero if j_μ is independent of A_μ as it was in the case treated in GGT. Incidentally, if we had been using the renormalized photon operators, the term in Eq. (2.6) involving $\partial D/\partial x_0$ would have made a contribution to Eq. (2.7).

[5] G. Källén, Helv. Phys. Acta **25**, 417 (1952); and Copenhagen Lectures, 1952 and 1953 (unpublished).

If now, following GGT, we restrict our attention to the case $x_0\to+\infty$, $y_0\to-\infty$, the second and third terms on the right hand side of Eq. (2.7) vanish and using Eq. (2.4) and the fact that $D(x)=-D(-x)$, we have

$$[A_\mu(x),A_\nu(y)]\to i\delta_{\mu\nu}D(x-y)$$
$$+\int d^4x'\int d^4y' D(x-x')D(y'-y)$$
$$\times\Big\{\eta(x'-y')[j_\mu(x'),j_\nu(y')]$$
$$-\delta(x_0'-y_0')\Big[j_\mu(x'),\frac{\partial A_\nu(y')}{\partial y_0'}\Big]\Big\}. \quad (2.8)$$

Needless to say, in practice the vanishing of the other terms in (2.7) requires some precise specification of how the limit $x_0\to+\infty$, $y_0\to-\infty$ is to be taken. We shall not discuss this point here.

It is useful to introduce into Eq. (2.8) the Fourier representation of the D function:

$$D(x)=-\frac{i}{(2\pi)^3}\int d^4k\epsilon(k)\delta(k^2)e^{ik\cdot x}. \quad (2.9)$$

We obtain then, finally,

$$[A_\mu(x),A_\nu(y)]\to i\delta_{\mu\nu}D(x-y)-i\int\frac{d^4k'}{(2\pi)^3}\int\frac{d^4k}{(2\pi)^3}$$
$$\times e^{ik'\cdot x-ik\cdot y}\epsilon(k')\epsilon(k)\delta(k^2)\delta(k'^2)\mathfrak{M}_{\mu\nu}(k',k), \quad (2.10)$$

where

$$\mathfrak{M}_{\mu\nu}(k',k)=-i\int d^4x'\int d^4y' e^{-ik\cdot x'+ik\cdot y'}$$
$$\times\Big\{\eta(x'-y')[j_\mu(x'),j_\nu(y')]$$
$$-\delta(x_0'-y_0')\Big[j_\mu(x'),\frac{\partial A_\nu(y')}{\partial y_0'}\Big]\Big\}. \quad (2.11)$$

The matrix element of this operator between an initial and final state of the matter system yields the generalization of the analogous quantity introduced in GGT, Eq. (3.10). In the next section we shall study the relation between such a matrix element and an exact expression for the photon scattering amplitude. The requirement that (2.10) shall vanish for space-like separations of x and y together with the relation between $\mathfrak{M}_{\mu\nu}(k',k)$ to the photon scattering amplitude to be discussed below leads, as is shown by GGT, to the Kramers-Kronig dispersion relations.

III. THE EXACT PHOTON SCATTERING AMPLITUDE

We wish to derive an expression for the scattering amplitude which describes the situation of a photon of four-momentum k and polarization ν being scattered by a matter system in state i, say, which makes a transition to state f producing a photon of four-momentum k',

polarization μ. For this purpose it is useful to consider the propagation function

$$G_+(x,y)=\langle f|P(A_\mu(x),A_\nu(y))|i\rangle. \quad (3.1)$$

It is easy to see that as $x_0\to+\infty$, $y_0\to-\infty$, (3.1) is proportional to a matrix element of the S-matrix, since $A_\mu(x)\to A_\mu^{\text{out}}(x)$ and $A_\nu(y)\to A_\nu^{\text{in}}(y)$ and $A_\mu^{\text{out}}(x)=S^{-1}A_\mu^{\text{in}}(x)S$.[5] We assume that the states i and f are steady. The matrix element of the R matrix ($S=1-iR$) which gives the scattering amplitude for the process described above may be found from the propagation function (3.1) in a familiar fashion.[6] The procedure given in reference 6 is equivalent to the following instruction:

$$\langle f,k',\mu|R|i,k,\nu\rangle=-i\int d^4x\int d^4y e^{-ik'\cdot x}$$
$$\times\{(-\Box_x^2)(-\Box_y^2)G_+(x,y)\}e^{ik\cdot y}. \quad (3.2)$$

We now carry out the differentiations indicated in (3.2). It follows from the canonical commutation rules and from the equation of motion for $A_\mu(x)$, Eq. (2.1), together with the definition of the P bracket that

$$-\Box_x^2G_+(x,y)=\langle f|P(j_\mu(x),A_\nu(y))|i\rangle$$
$$-i\delta(x-y)\delta_{fi}\delta_{\mu\nu}. \quad (3.3)$$

Let us assume that the states i and f are different so that we may ignore the second term in (3.3). Next we apply $-\Box_y^2$ and obtain

$$(-\Box_x^2)(-\Box_y^2)G_+(x,y)=\langle f|P(\dot{j}_\mu(x),j_\nu(y))$$
$$-\delta(x_0-y_0)[j_\mu(x),\partial A_\nu(y)/\partial y_0]|i\rangle. \quad (3.4)$$

In deriving Eq. (3.4) we have dropped a term involving $\delta(x_0-y_0)[A_\nu(y),j_\mu(x)]$ in accordance with our assumptions about $j_\mu(x)$. If we had been using renormalized photon operators, this term would make a contribution. It is, however, a C-number contribution if the only way time derivatives entered $j_\mu(x)$ were through the explicit charge-renormalization terms, and such contributions may again be dropped if the states i and f are different. The final expression for the scattering amplitude is then found to be

$$\langle f,k',\mu|R|i,k,\nu\rangle$$
$$=-i\int d^4x\int d^4y e^{-ik'\cdot x}\langle f|P(\dot{j}_\mu(x),j_\nu(y))$$
$$-\delta(x_0-y_0)[j_\mu(x),\partial A_\nu(y)/\partial y_0]|i\rangle e^{ik\cdot y}. \quad (3.5)$$

This looks exactly the same as the corresponding formula in GGT, Eq. (3.13) (remember that the commutator term should not appear in the case considered there), except for the fact that now the exact Heisenberg current operators appear in place of the interaction representation operators. Note that we have made no use of the fact that k and k' refer to real photons, i.e., $k^2=k'^2=0$, so that Eq. (3.5) is equally valid both on and off the energy shell.

A special case of Eq. (3.5) has previously been given by Low[7] in his treatment of the scattering of photons by nucleons interacting with charged mesons. Low's derivation is based on Dyson's expression for the S matrix in the interaction representation and will be published shortly by him. Similar results for special cases have also been given by Nambu.[8]

Our argument is now essentially completed. The matrix element appearing in Eq. (3.5) is the same as that which is found by forming $\langle f|\mathfrak{M}_{\mu\nu}(k',k)|i\rangle$ from Eq. (2.11) except that the P-bracket of the current operators appears instead of the commutator times a step function. This is just the situation found in the perturbation treatment of GGT, and the same relations among the dispersive and absorptive parts of the on-energy-shell matrix elements for forward scattering ($k'=k$) given there [GGT, Eqs. (3.15) and (3.16)] now hold as rigorous relations. The discussion given by them about the analytic properties of the scattering amplitudes may be taken over directly.

IV. CONCLUSIONS

It is, of course, not surprising that the Kramers-Kronig relations which follow from the above results are correct, independent of perturbation theory, and under very general assumptions on the form of the coupling between the electromagnetic field and the matter field. It is gratifying, however, that such a simple derivation may be given.

The rigorous expression for the scattering amplitude derived in Sec. III appears to be quite useful. An exactly analogous expression may be derived for the problem of meson-nucleon scattering. From such expressions many general properties of the scattering amplitudes may be deduced in a very simple way. These together with other applications will be discussed in a subsequent publication.

Note added in proof.—After this manuscript had been submitted for publication, the author received a letter from Dr. S. S. Schweber outlining a very similar treatment of the problem discussed here which he had independently given at about the same time.

Dr. Schweber's derivation differs from ours in the following respects: (1) He uses both "in" and "out" fields, expressing [see Eqs. (2.2) and (2.5)] $A_\mu(x)$ in terms of $A_\mu^{\text{in}}(x)$ and the retarded Green's function and $A_\nu(y)$ in terms of $A_\nu^{\text{out}}(y)$ and the advanced Green's function. (2) By limiting himself to a consideration of $\langle f|[A_\mu(x),A_\nu(y)]|i\rangle$ in the limit as $x_0\to+\infty$ and $y_0\to-\infty$, he shows that the use of Eq. (3.6) may be avoided, at least in the case where j_μ is independent of A_μ, which is the only case treated by him.

[6] Karplus, Kivelson, and Martin, Phys. Rev. **90**, 1073 (1953).

[7] F. E. Low, Phys. Rev. **96**, 1428 (1954). The author wishes to thank Prof. Low for a preprint of his paper and for a discussion of his derivation of Eq. (3.2) for his problem.

[8] Y. Nambu (private communication).

PHYSICAL REVIEW VOLUME 97, NUMBER 2 JANUARY 15, 1955

Geons*

JOHN ARCHIBALD WHEELER
Palmer Physical Laboratory, Princeton University, Princeton, New Jersey
(Received September 8, 1954)

Associated with an electromagnetic disturbance is a mass, the gravitational attraction of which under appropriate circumstances is capable of holding the disturbance together for a time long in comparison with the characteristic periods of the system. Such gravitational-electromagnetic entities, or "geons"; are analyzed via classical relativity theory. They furnish for the first time a completely classical, divergence-free, self-consistent picture of the Newtonian concept of body over the range of masses from $\sim 10^{39}$ g to $\sim 10^{67}$ g. Smaller geons are quantum objects whose analysis would call for the treatment of characteristic new effects. Topics covered in the discussion include: 1. Need for a self-consistent formulation of the concept of "body" in classical physics; geons *vs* free waves; electrical neutrality of geon; size and mass relations; the quantum limit and electron pair phenomena. 2. Orders of magnitude for toroidal geons; first estimates of leakage rates; a "phosphor" model of a geon; attrition and attritivity; energy action relation. 3. Idealized spherical geon; conditions required for symmetry; instability relative to pairing of light rays; time scale of instability long compared to vibration periods; spherical metric; wave equation for electromagnetic potential; evaluation of stress-energy tensor; its position as source of gravitation field; the gravitational field equations; the three equations of the self-consistent geon; simplification by scale transformation; first analysis of the eigenvalue problem; further scale transformation to get behavior of solution in active region of geon; further analysis of eigenvalue dependence; electronic calculator integration of equations of self-consistent geon; mass and radius values. 4. Transformations and interactions of electromagnetic geons; evaluation of refractive index barrier penetration integral for spherical geon; photon-photon collision processes as additional mechanism for escape of energy from system; restatement in language of coupling of characteristic modes; the thermal geon; comparison of gravitation and virtual electron pair phenomena as sources of coupling between modes; gravitational coupling and collective vibrations of geon; fission of a geon; interaction between two geons simple at large distances; orientation dependence and exponential term at intermediate distances; violent transmutation processes in closer encounters. 5. Influence of virtual pairs on geon structure; description in terms of refractive index correction; relation to photon-photon collision picture; more precise formulation via Heisenberg-Euler electrodynamics; corrections to stress-energy tensor and electromagnetic field equations. 6. Neutrino-containing geons; general similarity to electromagnetic geons; specificity of geon-geon interactions; the size subject to simple analysis unexpectedly limited by neutrino-neutrino encounters and the process $\nu+\nu\rightarrow\mu+e$; similarity of size limitation to that for electromagnetic geons; comments on present status of neutrino theory of light. 7. Electricity, Gauss's theorem, and gravitational field fluctuations. 8. Conclusions: the geon completes the scheme of classical physics; one's interest in following geons into quantum domain will depend upon one's view of the relation between very small geons and elementary particles.

1. INTRODUCTION AND SUMMARY

THE position of the concept of "body"[1] in the general theory of relativity[2] has always been interesting. A planet moving as a body along a geodesic is the idealization behind one of the most important predictions of the general theory, and certainly the one that has received the most thorough observational con-

* The word "geon" is used here as an abbreviation for the phrase "gravitational-electromagnetic entity" and in place of the name kugelblitz or ball of light previously used in a survey of the problems of fields and particles: Am. J. Phys. (to be published); for some other aspects of these problems, see also J. phys. rad. (to be published); see also the point of view ascribed by the author to Sugawara-no-Michizane in Proc. Phys. Soc. Japan 9, 36 (1954). The present article is part V of a study of classical field theory. I is unwritten. In parts II, J. A. Wheeler and R. P. Feynman, Revs. Modern Phys. 17, 157 (1945), III, J. A. Wheeler and R. P. Feynman, Revs. Modern Phys. 21, 425 (1949), and IV, in preparation in collaboration with Professor Gilbert N. Plass, the emphasis is upon action at a distance, not as a way of understanding charged bodies, but as a way of understanding the fields that act between them. In the present paper the emphasis is on the fields interior to a *classical* body.

[1] The noun "body" is used here to connote an object possesed of a mass and of three position coordinates, and subject both externally and internally to a classical description, in contrast to the notion of a particle, with at least the interior of which quantum properties are associated.

[2] By "general theory of relativity" is meant here that battle-tested system of ideas and equations which Einstein developed in 1915 to describe gravitation and electromagnetism. I am indebted to Professor Einstein for several interesting discussions of the evolution of his ideas and of their relation to Newtonian concepts. Excluded here from the phrase "general theory" is the cosmological term, absent with good reason from the original formulation, introduced only when it was found that the original theory could not account for a static universe, and disowned when the universe was found not to be static. Excluded also are varied modifications of general relativity which attempt to "unify" gravitation and electricity, the latest of which has been shown by Joseph Callaway, Phys. Rev. **92**, 1567 (1953) to predict that a test particle will move as if uncharged, no matter how much charge is loaded upon it. In the accepted formulation of general relativity—to use the metric conventions of Pauli's treatise—the proper distance, ds, or proper interval of cotime (c times time), $d\tau$, between two neighboring events is given by

$$(ds)^2=-(d\tau)^2=g_{\alpha\beta}dx^\alpha dx^\beta.$$

The state of the space time continuum is specified by giving in addition to the ten gravitational potentials, $g_{ik}=g_{ki}$, the four electromagnetic potentials, A_i. The electromagnetic field, $F_{ik}=\partial A_k/\partial x^i-\partial A_i/\partial x^k$, has six distinct components, whereas gravitational effects are expressed in a covariant way by the twenty distinct components of the curvature tensor R_{abcd}; or by the mixed components of the same tensor,

$$R^h{}_{ikj}=\partial\Gamma_{ij}{}^h/\partial x^k-\partial\Gamma_{ik}{}^h/\partial x^j+\Gamma_{kd}{}^h\Gamma_{ij}{}^d-\Gamma_{jd}{}^h\Gamma_{ik}{}^d,$$

where the typical Γ is an abbreviation for

$$\Gamma_{ij,k}=\tfrac{1}{2}(\partial g_{jk}/\partial x^i+\partial g_{ik}/\partial x^j-\partial g_{ij}/\partial x^k).$$

Of the twenty distinct components of R_{abcd}, ten remain in the contracted curvature tensor $R_{ik}=R_{ia}{}^{k}{}^{\alpha}$; and of these one remains in the curvature invariant $R=g^{\alpha\beta}R_{\alpha\beta}$. The Einstein tensor $G_{ik}=R_{ik}-\frac{1}{2}g_{ik}R$ may be considered to be the translation into the language of general relativity of the notion of d'Alembertian of the gravitational potentials, g_{ik}. The gravitational field equations thus have the form

$$G_{ik}=(8\pi G/c^4)T_{ik},$$

where $G=6.67\times10^{-8}$ cm^3/g sec^2 is the Newtonian constant of gravitation, $c=3.00\times10^{10}$ cm/sec is the speed of light, and T_{ik} is the symmetric stress-momentum-energy tensor of the electro-

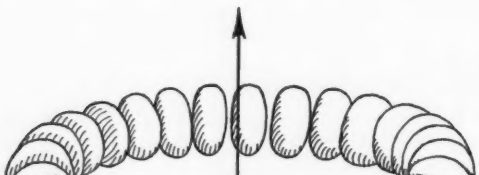


FIG. 1. Regions of strong electric field strength in a simple toroidal geon of zero angular momentum. Two waves of equal strength run around the torus in opposite directions to produce a standing wave, with electric fields strong in the regions indicated, and magnetic fields strong in the region between. The gravitational field created by this disturbance, and required to hold the disturbance together, is representable to a good approximation as static and independent of azimuth.

firmation.[3] In recent years a great advance has been made in the theory. The geodesic equation that determines the motion of the object in a known field, and the field equations that find the metric from the motions of the masses, have been shown not both to be necessary, as they have been in every other formulation of physics. Instead, the equation of motion of the body has been shown to follow as a *consequence* of the field equations.[4] This circumstance, the fact that relativity is the first description of nature that makes geometry a part of physics, the absence of any acceptable alternative of comparable scope, and the battle-tested internal consistency of the theory, all make it is necessary to take seriously general relativity and explore further its consequences and concepts. Of these concepts the notion of body is in an unhappy state. Either one sticks to general relativity as it is, and treats the object as a singularity in the metric, or one postulates that the field is regular everywhere, and counts on quantum theory somehow to explain how this can be so even in the region of the body. Both approaches lead for the time being to an impasse. For this reason it is interesting to discover that there exists a third possibility. On the basis of classical general relativity as it already exists, and without any call on quantum theory, it turns out to be possible to construct an entity that we call geon. This object serves as a classical, singularity free, exemplar of the "bodies" of classical physics. Of such entities there exist in principle a great variety, distinguished from one another by mass, intrinsic angular momentum, and other properties. The simplest variety (Fig. 1) is most easily visualized as a standing electromagnetic wave, or beam of light, bent into a closed circular toroid of high energy concentration. It is held in this form by the gravitational attraction of the mass associated with the field energy itself. It is a self-consistent solution of the problem of coupled electromagnetic and gravitational fields, as defined by the system of equations,

$$G_{ik}=(8\pi G/c^4)T_{ik}, \tag{1}$$

$$T_i{}^k=(1/4\pi)(F_{i\alpha}F^{k\alpha}-\tfrac{1}{4}F_{\alpha\beta}F^{\alpha\beta}\delta_i{}^k), \tag{2}$$

and

$$(-g)^{-\frac{1}{2}}(\partial/\partial x^\alpha)(-g)^{\frac{1}{2}}F^{i\alpha}=0, \tag{3}$$

all derivable from the action principle:

$$\delta\int\int\int\int\{(1/16\pi c)F_{\alpha\beta}F^{\alpha\beta} - (c^3/16\pi G)R\}d(\text{four-volume})=0. \tag{4}$$

An order of magnitude discussion of the properties of these entities appears in Sec. 2. Section 3 presents theory and results of electronic machine solutions of the self-consistent field equations for the particularly simple but rather idealized case of a spherical geon. The existence of geons would seem to impart to classical general relativity theory a comprehensiveness for which one had not dared to hope. This theory turns out not only to account for the fields produced by bodies, and the motions of bodies, but even to explain why there should be bodies. In this sense classical relativity theory would seem to be revealed as a logically self-consistent and completed whole of unexpected comprehensiveness.

Having geons as model for the bodies of classical physics, we can put into a new perspective some parts of general relativity. First, the geodesic equation of motion of a body displays itself clearly as an idealization. To be able to give any meaning to the metric that appears in this equation, we have to be able to speak of the "background field"—the gravitational and electromagnetic magnitudes that would be present in the absence of the geon. This will only be possible when the total fields at some distance from the object vary slowly over a distance comparable to the linear exten-

---

magnetic field,

$$T_{ik}=(\partial/\partial g^{ik})(1/8\pi)F_{\alpha\beta}F^{\alpha\beta}.$$

Of the equations for the electromagnetic field itself, half are already automatically satisfied by virtue of the way the fields are expressed in terms of the potentials, A_i. The other half are usually written as inhomogeneous equations,

$$(-g)^{-\frac{1}{2}}(\partial/\partial x^\alpha)(-g)^{\frac{1}{2}}F^{i\alpha}=4\pi s^i,$$

where g is the determinant of the g_{ik}, or where $(-g)^{\frac{1}{2}}$ is the ratio of an element of four volume to the product $dx^1dx^2dx^3dx^4$, and where the four vector s^i describes the density and flow of free electricity. We omit this source term (a) because we do not need it in the considerations of this article (b) because there is no self-consistent classical theory for the existence of electric charge (c) because the considerations of Sec. 1 and Sec. 7 suggest that the appearance of free electricity is a quantum phenomenon. Thus limited to charge free space, general relativity constitutes a well defined, self-consistent, self-contained whole.

[3] G. M. Clemence, Revs. Modern Phys. **19**, 361 (1947); Proc. Am. Phil. Soc. **93**, 532 (1949).

[4] L. Infeld and A. Schild, Revs. Modern Phys. **21**, 408 (1949); see also Einstein, Infeld, and Hoffmann, Ann. Math. **39**, 65 (1938); D. M. Chase, Phys. Rev. **95**, 243 (1954); and a paper by L. Infeld, Acta Phys. Polonica (to be published) for a perusal and discussion of which I am indebted to Professor Infeld.

sion of the geon. In this case the object will accelerate coherently under the action of external forces. The happenings fit into the scheme of physics envisaged by Newton and Maxwell: preexisting bodies generating forces and being acted upon by forces. The appropriate modifications to pre-Einsteinian physics associated with special and general relativity of course add themselves automatically. On the other hand, when the fields outside the geon approach non-uniform values, then this entity does not react as a unit. Its internal degrees of freedom are disturbed. When the external inhomogeneities are strong, the system will break into two or more parts. There is nothing in what we know today of physics that makes unreasonable the failure under such conditions of the idealization of body. Evidently there exist conditions where the idealization of "body" loses all significance, and where there comes into play a new kind of physics of geons—but a physics that is still entirely classical (Sec. 4).

The existence of geon transformation processes makes clear a second point about classical relativity physics, that there exists in principle no sharp distinction between geons as concentrations of electromagnetic energy, capable of break-up and integration processes, and the "free" electromagnetic waves that pass through the space between geons, and experience scattering, absorption, or emission by geons. Legalistically speaking, the state of the universe of classical physics is described by the singularity free electromagnetic and gravitational magnitudes at every point, and by nothing more.

Finally, physics as rounded out by the geon concept forms the only fully comprehensive and self-consistent system of classical physics that we possess. This picture of bodies and fields leaves as little place for free electric charges as for isolated magnetic poles: none at all. The electromagnetic field is as free of singularity as the gravitational field. Lines of forces never begin and never end. At most they form loops that shrink to zero and disappear, or reappear, as time goes on and field strengths change. No natural way is evident to escape the conclusion that the quantum of electricity has to do with the quantum of action and does not belong within the framework of a *completely* classical physics. Considering that our physical world extends at least from 10^{-13} cm to 10^{27} cm, and that in the face of these forty decades or more e^2 differs from $\hbar c$ by only two powers of ten, it would also seem that one has some independent reason to believe that free electricity is a quantum phenomenon. Therefore it appears reasonable to accept as a description of a *classical* body the always electrically neutral geon.

What is the range of sizes of classical geons? These objects obey a simple scaling law. From one self consistent solution of the coupled equations of gravitation and electromagnetism there follows another solution of of $(1/n)$th the mass, by decreasing all distances by a factor n, increasing all electromagnetic field strengths by a factor n, and leaving all the gravitational potentials g_{ik} unaltered in strength. In this change the classical action integral, J, associated with the field disturbance falls to $1/n^2$ of its original value. Such a scaling law must exist because the field equations contain only the gravitation constant and the velocity of light, out of which one can form no quantity with the dimensions either of a length or a mass or a field strength or an action. Consistent with the scaling law, but even more specific, are the order-of-magnitude formulas of Sec. 2 for the simplest circular toroidal geons, with a number of standing waves around the circumference equal to an integer or azimuthal index number, a:

$$\begin{aligned}
\text{mass} &\sim a^{\frac{1}{2}}\ (\text{action}\cdot c/G)^{\frac{1}{2}},\\
\text{major radius} &\sim a^{\frac{1}{2}}\ (\text{action}\cdot G/c^3)^{\frac{1}{2}},\\
\text{minor radius} &\sim \text{small multiple of } \lambda\!\!\!^{-},\\
\lambda\!\!\!^{-} = (\text{wavelength}/2\pi) &\sim a^{-\frac{1}{2}}\ (\text{action}\cdot G/c^3)^{\frac{1}{2}},\\
\text{frequency} &\sim a^{\frac{1}{2}}\ (c^5/G\cdot\text{action})^{\frac{1}{2}},\\
\left\{\begin{matrix}\text{mass density in}\\ \text{active region}\end{matrix}\right\} &\sim \left\{\begin{matrix}\text{reasonable fraction of}\\ (ac^5/G^2\cdot\text{action})\end{matrix}\right\},\\
\left\{\begin{matrix}\text{peak field values}\\ \text{in electrostatic}\\ \text{volts/cm or gauss}\end{matrix}\right\} &\sim \left\{\begin{matrix}\text{reasonable fraction of}\\ (ac^7/G^2\cdot\text{action})^{\frac{1}{2}}\end{matrix}\right\}. \qquad (5)
\end{aligned}$$

The upper limit to the size of geons is the linear extension of the universe itself. This limit shows itself in the following way. The self-consistent field equations of the geon possess solutions, the properties of which depend upon the boundary conditions. In simple geon theory we impose the requirement that g_{ik}'s are asymptotic to a flat metric at distances large in comparison with the extension of the geon. This boundary condition has to be modified when the size of the geon is comparable to the radius of curvature of a closed universe. Then the relations (5) undergo characteristic corrections, such that the mass of the geon can never exceed the mass of the universe.

Incidentally it is interesting to notice that a purely classical closed universe consisting of a large number, N, of geons of comparable size must consist mostly of empty space. This conclusion follows from the fact that the proportionality factor between radius and mass,

$$\text{radius} \sim (G/c^2)\ \text{mass} = (0.74\times 10^{-28}\ \text{cm/g})\ \text{mass}$$

is the same in order of magnitude for geons as for the universe itself. Consequently a ratio $1/N$ between the mass of the parts and the mass of the whole imples a ratio $\sim N(1/N)^3 = 1/N^2$ between the volume of the parts and the volume of the whole.

To find the lower limit to the size of classical geons, we have to investigate in turn the several physical magnitudes that characterize this object, and see which of these magnitudes, with decreasing geon size, first passes out of the domain of application of classical theory. (1) Is action the critical magnitude? Certainly

one will be in the quantum domain when the action is comparable to the quantum of angular momentum, $\hbar = 1.054 \times 10^{-27}$ g cm^2/sec. Inserting this value into the formulas (5), we find

$$\text{mass} \sim a^{\frac{1}{2}} 2.18 \times 10^{-5} \text{ g},$$

$$\text{major radius} \sim a^{\frac{1}{2}} 1.63 \times 10^{-33} \text{ cm}.$$

We can stop here with the evaluation. We are evidently far below the limit where it might have been right to use classical theory, and *quite outside the domain of application* of Eqs. (5). Electron theory allows no possibility to deal with distances smaller than the reduced Compton wavelength, $\hbar/mc = 3.87 \times 10^{-11}$ cm, without taking into account complex and specifically quantum mechanical fluctuations in the distribution of pairs and in the electromagnetic field. What about fluctuations in the gravitational field? They appear to be negligible on any ordinary scale of distances, as one sees from the following reasoning. In a region with space and cotime extensions of the order of L the relevant fluctuations in the electromagnetic field[5] are of the order $\Delta \mathcal{E} \sim (\hbar c)^{\frac{1}{2}}/L^2$, and $\Delta A \sim (\hbar c)^{\frac{1}{2}}/L$ in the potentials. The similarity in character of the gravitational and electromagnetic field equations, and the difference in the way they contain the various fundamental constants, indicate that the corresponding formula for the fluctuation in a typical gravitational potential is

$$\Delta g \sim (\hbar G/c^3)^{\frac{1}{2}}/L. \tag{6}$$

These fluctuations will be inappreciable in comparison to typical average values of the metric, $g \sim 1$, so long as the distances, L, under consideration are substantial in comparison with the characteristic length,

$$L^* = (\hbar G/c^3)^{\frac{1}{2}} = 1.63 \times 10^{-33} \text{ cm}. \tag{7}$$

This is exactly the distance in comparison with which we have just concluded that classical geons must be enormous.

(2) Is distance the critical magnitude? We set the reduced wavelength in Eqs. (5) equal to $\hbar/mc$ and solve for the action and other physical magnitudes, finding

$$\text{action} \sim a(\hbar c/m^2 G)\hbar = a 5.72 \times 10^{44} \hbar,$$

$$\text{mass} \sim a(5.69 \times 10^{44})^{\frac{1}{2}} 2.15 \times 10^{-5} \text{ g} = a 5.12 \times 10^{17} \text{ g},$$

$$\text{major radius} \sim a 3.87 \times 10^{-11} \text{ cm},$$

$$\text{minor radius} \sim \text{small multiple of } 3.87 \times 10^{-11} \text{ cm},$$

$$(\text{wavelength}/2\pi) \sim 3.87 \times 10^{-11} \text{ cm},$$

$$\text{frequency} \sim mc^2/\hbar,$$

all of which is so far acceptable; but the mass density in the active region,

$$\text{geon density} \sim \text{a reasonable fraction of } (c^2/G)(mc/\hbar)^{\frac{1}{2}} = 0.90 \times 10^{49} \text{ g/cm}^3$$

is obviously fantastically in excess of the density of nuclear matter,

$$\text{nuclear density} \sim 1836m/[(4\pi/3)(e^2/2mc^2)^3] = 1.4 \times 10^{14} \text{ g/cm}^3. \tag{8}$$

Consequently we still find ourselves in a domain where the classical geon relations (5) *cannot be applied.*

(3) Is density the limiting magnitude? Taking this time nuclear densities, (8), as the reference point, we find from (5) values of the action, mass, radius, and frequency of the geon which are safely outside the obvious quantum limits; but the field strength in the active region,

$$(c^2 1.4 \times 10^{14} \text{ g/cm}^3)^{\frac{1}{2}} = 3.6 \times 10^{17} \text{ (gauss or electrostatic volts/cm)},$$

is then 8000 times greater than the critical value met in the quantum electrodynamics of the vacuum,

$$\mathcal{E}_{\text{crit}} = m^2 c^3/e\hbar = 4.42 \times 10^{13} \text{ (gauss or electrostatic volts/cm)}. \tag{9}$$

Thus, when an electric field $\mathcal{E}$, working on an elementary charge e over the localizability distance $\hbar/mc$, can impart to an electron an energy of the order mc^2, then this electric field will bring forth from empty space pairs of positive and negative electrons. Under such conditions the geon requires for its description the specification of the state of the pairs as well as the statement of the electromagnetic and gravitational field strengths. Again we find ourselves outside the range of validity of the classical geon equations (5).

(4) We conclude that the *critical field strength,* $\mathcal{E}_{\text{crit}}$, *of pair theory marks the lower limit of classical geons.* Thus, simple *toroidal* electromagnetic geons are only then purely classical entities, when their magnitudes are on the large geon side of the following limits:

$$\text{action} \sim \text{reasonable fraction of } ac^7/G^2\mathcal{E}_{\text{crit}}^2 = a\hbar(e^2/Gm^2)^2(\hbar e/e^2) = 2.38 \times 10^{87} a\hbar,$$

$$\text{mass} \sim \text{reasonable fraction of } ac^4/G^{\frac{1}{2}}\mathcal{E}_{\text{crit}} = a 1.065 \times 10^{39} \text{ g},$$

$$\text{major radius} \sim \text{reasonable fraction of } ac^2/G^{\frac{1}{2}}\mathcal{E}_{\text{crit}} = a 0.791 \times 10^{11} \text{ cm},$$

$$\text{minor radius} \sim 0.791 \times 10^{11} \text{ cm},$$

$$(\text{wavelength}/2\pi) \sim 0.791 \times 10^{11} \text{ cm},$$

$$\text{frequency} \sim 1 \text{ vibration}/16.6 \text{ sec},$$

$$\text{mass density} \sim \mathcal{E}_{\text{crit}}^2/c^2 = 2.16 \times 10^6 \text{ g/cm}^3,$$

$$\text{peak electric field} \sim \mathcal{E}_{\text{crit}} \text{ of Eq. (9)}. \tag{10}$$

One has only to compare these critical dimensions with the properties of the sun, mass $= 1.97 \times 10^{33}$ g, radius $= 6.95 \times 10^{10}$ cm, to conclude that even the lightest *classical* geons form entities enormous in comparison to the objects studied in the laboratory. But the critical

[5] N. Bohr and L. Rosenfeld, Kgl. Danske Videnskab. Selskab., Mat-fys. Medd. 12, No. 8 (1933); Phys. Rev. 78, 794 (1950).

dimensions are still small compared to the scale of the universe. There is ample room between 10^{11} cm and 10^{28} cm to talk of classical geons and their motions, interactions and transformations. In this sense we continue to regard classical general relativity as a completed self contained subject with a well-defined scope of its own.

Having surveyed the boundaries of the classical theory of geons, we can touch on a few of the implications of quantum theory for these objects. First, as to the general situation, it is clear that the critical magnitudes of Eq. (10) prevent us in no way whatever from considering geons of lower mass. It is only required that we take quantum effects properly into account. It is also clear that these effects will be of various kinds. As we move down in mass, first one effect, previously unimportant, will become decisive, then another effect, and so on. In each region certain idealizations will be appropriate. In each region the formulas connecting mass and other geon properties with the quantity of action will have characteristic forms of their own. Presumably the investigation of each region will present successively greater difficulties. Second, the quantum region earliest encountered (Sec. 5) will lie between the realm of classical electrodynamics and the domain where fields are so strong that substantial numbers of pairs appear. In this intermediate "Region II," the fields vary in space and time fantastically slowly in comparison with the characteristic distances and times of electron theory. Moreover, they are strong enough to produce only virtual pairs. These pairs give rise to charges and currents that make a substantial contribution to the total field, as first analyzed in detail by Euler and Kockel and Heisenberg.[6] In quantum language, if two opposed photons of high frequency can produce a pair, then two opposed photons of low frequency can produce a virtual pair. With the reannihilation of the pair the photons go off in altered directions; hence a scattering of light by light[7] and an effectively nonlinear electrodynamics. The consequence of this nonlinearity for a simple toroidal electromagnetic geon is most simply envisaged as an increase of refractive index in the region of concentration of electromagnetic energy. The effective value of the speed of light is reduced. Geon mass no longer scales in proportion to the square root of the action variable. Instead, as the action of the system diminishes, the mass appears to fall faster than linearly [Eq. (81)]. However, soon we are into Region III, where real pairs first appear in large numbers. Here new studies must be made before any results can be stated.

As third circumstance in the relation of the quantum to the geon, there exist neutrinos: entities apparently coordinate with photons in the description of nature. Their Fermi-Dirac statistics, unlike those of photons, makes it impossible for more than one to be accommodated in a state of definite wave number and polarization. From them no disturbance can be built up of a classical magnitude, describable in correspondence principle terms. They have an inescapably quantum character. Apart from this circumstance they can be used in the construction of geons just as well as photons. In addition to purely electromagnetic geons there consequently also exist in principle (Sec. 6) purely neutrino geons, and geons of mixed type. Thus half integral as well as integral spins are permitted to geons.

As last implication of the quantum for geon physics, the fluctuations in the gravitational metric are inescapable, little as one can say (Sec. 7) about their consequences for the validity of Gausses theorem and for the existence of free electric charge.

Details of the discussion follow.

2. ORDERS OF MAGNITUDE FOR SIMPLE TOROIDAL ELECTROMAGNETIC GEONS

The gravitational deflection of a pencil of light into a torus is no different in principle from the deflection of light by the sun. Apart from factors of the order of two[8] one can estimate the deflection by equating a kinematic acceleration to a gravitational acceleration:

$$c^2/R \sim GM/R^2.$$

Thus the radius R of the torus and the mass M are connected by the relation

$$R \sim (G/c^2)M = (0.741 \times 10^{-28}\ \text{cm/g})M, \qquad (11)$$

This formula is familiar in another connection, for it supplies the well known relation between the mass of an object and its so called gravitational radius, a measure of the distances at which the gravitational potentials depart significantly from the values appropriate to flat space. This circumstance makes it evident that any accurate treatment of geon properties has to be carried out within the framework of general relativity.

Near a section of the torus the gravitational field will resemble approximately the field due to an infinitely long cylindrically symmetrical distribution of mass. It will increase inversely as the first power of the distance, ρ, from the center of the cylinder, until the point of observation moves into the region of strong energy concentration. There the gravitational potential will no longer continue its logarithmic increase.

The narrowness of the pencil of light, or of the region of strong energy concentration, will depend upon the wavelength. Define by a large integer a an azimuthal index number, and let the wavelength constitute the small fraction, $1/a$, of the circumference:

$$(\text{wavelength}/2\pi) \equiv \lambda\!\!\!^{-} \sim R/a.$$

[6] H. Euler and B. Kockel, Naturwiss. 23, 246 (1935); W. Heisenberg and H. Euler, Physik 98, 714 (1936).

[7] Low-frequency cross section in reference 6; high-frequency cross section in A. Achieser, Physik Z. Sowjetunion 11, 263 (1937); forward scattering at all frequencies calculated by R. Karplus and M. Neuman, Phys. Rev. 83, 776 (1951); derived from dispersion relation by J. Toll (to be published).

[8] See for example R. Tolman, *Relativity, Thermodynamics, and Cosmology* (Oxford University Press, Clarendon Press, 1934).

The pencil of light cannot be concentrated into a region smaller in lateral extension than the order of magnitude of $\lambda\!\!\!^{-}$ by any type of variation of refractive index with distance. Owing to the logarithmic variation of gravitational potential and effective refractive index near the cylinder, the lateral extension of the disturbance in the case of the torus will be a small multiple of $\lambda\!\!\!^{-}$, say $F\lambda\!\!\!^{-}$, where the dimensionless factor F will be approximately represented by a simple power of $\ln(R/\lambda\!\!\!^{-})=\ln a$. Even if a equals 10^{10}, this logarithm is only 23.

The circumstance that the minor radius of the torus is significantly larger than $\lambda\!\!\!^{-}$ implies that the gravitation field can be taken to be static in a reasonable approximation. Thus the source of the gravitation field, in accordance with the principle of equivalence of mass and energy, is the electromagnetic stress-energy tensor. This tensor varies rapidly over a distance of the order of $\lambda\!\!\!^{-}$. But the elementary contribution to the gravitation field is essentially a $1/r^2$, or long-range, field. The field, even in the most active part of the geon, comes mostly from distances of the order of the minor radius, $F\lambda\!\!\!^{-}$. Consequently local variations in gravitational source strength are relatively unimportant. What does count is the source strength averaged over a wavelength. In other words, the gravitational field may be regarded as depending upon ρ and z, to use cylindrical polar coordinates, but not on φ and t.

Of course nothing prevents consideration of toroidal geons of small azimuthal index, a, but then it is no longer such a good approximation to treat the gravitational field as static, and the situation is much more complicated to discuss. Moreover, such geons disintegrate rapidly.

Even for the case of large azimuthal index number there is a difference in simplicity between those toroidal geons that carry equal electromagnetic waves going in the positive and negative senses around the ring, and those where the disturbance runs all one way, or there are two disturbances of different magnitudes. In the first case we have to do with a system of zero angular momentum. The electromagnetic disturbance is a simple standing wave. Many components of the metric tensor vanish when expressed in cylindrical polar coordinates.[9] Significantly more distinct components of g_{ik} have to be considered when the system has angular momentum.[10] Particularly interesting is the case when the disturbance is unidirectional. Then the angular momentum of the system is of the order

$$\text{angular momentum}\sim McR\sim(c^3/G)R^2. \tag{13}$$

In the active region of a simple classical toroidal electromagnetic geon the field strengths will be of the order of magnitude $\mathcal{E}$, where

$$\mathcal{E}^2R(\Delta R)^2\sim Mc^2. \tag{14}$$

Consequently $\mathcal{E}\Delta R$, the order of magnitude of the typical difference of potential between two sides of the pencil of radiation, will have a universal value,

$$\text{potential difference}\sim(Mc^2/R)^{\frac{1}{2}}\sim c^2/G^{\frac{1}{2}}$$
$$=3.49\times10^{24}(\text{gauss cm or electrostatic volts}), \tag{15}$$

enough to impart to a particle of electronic charge and mass an energy greater than its rest energy by a factor $(e^2/Gm^2)^{\frac{1}{2}}$. This factor is the root of the ratio of the electric and gravitational forces between two electrons, and has the value $(4.17\times10^{42})^{\frac{1}{2}}=2.04\times10^{21}$.

Through the symmetry axis of the torus pass a typical meridian plane, and note the point in this plane where it intersects the region of maximum field strength in the torus. Then over a circle of radius $\sim\Delta R$ centered on this point the field strength has values smaller than the peak magnitude by a factor of only one or two powers of two. At distances from the most active point several times the magnitude ΔR, the field strengths fall off exponentially with a characteristic decay length of the order of magnitude of $\lambda\!\!\!^{-}$. This type of decay is familiar from the study of the propagation of light along the length of a long thin solid glass rod. The disturbance in the space surrounding the rod also has a characteristic decrement length of the order of $\lambda\!\!\!^{-}$. In the case of the geon we deal with a medium whose effective refractive index is nonuniform. The gravitational field has fallen off substantially at distances from the symmetry center of the geon of the order of magnitude of $2R$ and greater. Consequently the electromagnetic disturbance sufficiently far from the geon finds itself once again in an allowed region, where it can propagate normally. The strength of the electromagnetic field outside is an exceedingly small fraction of the strength inside, the approximate value of this fraction being given by an algebraic function of the large number $R/\lambda\!\!\!^{-}=a$ multiplied by an approximately exponential factor of the form $e(-\text{constant}\cdot a)$.

The geon is thus not in principle an isolated entity. The object in question, and every other classical geon in a classical universe, are nothing but manifestations of the same all pervading electromagnetic field. But the field outside is so extremely small in comparison with the field inside that for most purposes the geon has the character of a well-defined body.

Whether the field outside the geon is an outgoing wave or an incoming wave or a standing wave depends upon the initial conditions. It will be most relevant to consider here, as in most problems of physics, an irreversible dissipation of energy. The external wave then has no incoming component. It describes a continual transport of energy—and mass—out of the geon. We are faced with a purely classical analog to the Gamow-Condon-Gurney theory of radioactive alpha decay.

[9] Frederick J. Ernst, Jr., "Cylindrically Symmetric Fields in General Relativity," Junior Independent Paper, Princeton, 5 May 1954 (unpublished).

[10] For examples of gravitational sources endowed with angular momentum, see for example, G. E. Tauber, Ph.D. thesis, University of Minnesota, 1952 (unpublished).

It seems appropriate to give the same name of radioactive decay to the geon process, though for our present purposes this process is to be regarded as having absolutely no quantum character. Of course, were we temporarily to abandon the ground of classical physics, we would describe the emission process in the language of quanta.

As a model showing the slow decay characteristic of a relatively stable geon, one can consider a sphere of highly transparent glass with a radius of the order of ten wavelengths of light. A high-speed electron which enters the sphere produces ions and excites electromagnetic disturbances including modes of vibration of visible wavelength. Of many of these modes the energy will leak out of the refractive index barrier quickly because of the nearly normal incidence of the equivalent ray system upon the air interface; but those described by spherical harmonics of the highest order, and therefore represented by the most nearly tangential rays, will be endowed with lives as long as 10^{-6} sec. The exponential fall of the relevant modes in the air just outside the sphere "phosphor" will be similar to the behavior in the air above a totally internally reflecting prism. In that case the decay continues indefinitely with distance; but in the spherical case the decay ceases, and free oscillation commences, at a distance comparable to the radius of curvature of the surface.

As the geon slowly loses mass, it shrinks in accordance with the similarity law for these bodies. At the same time the circular frequency, ω, of the emerging electromagnetic disturbances goes up in inverse proportion to the radius and mass. Consequently the ratio of the loss per unit time to what remains at that time is not a constant, as in the usual theory of radioactive decay, but is proportional to the current frequency scale of geon processes:

$$dM/M = dR/R = -d\omega/\omega = -\alpha\omega dt. \tag{16}$$

Here α is a dimensionless constant, the "attrition," or fractional loss of mass per radian of the electromagnetic vibration. Integration of (16) gives

$$-t = 1/\alpha\omega; \tag{17}$$

the mass of the geon decreases linearly with time. The negative sign distinguishes the time at which the geon is observed to have the frequency ω from the time at which it collapses (quantum modifications in the later stages of the decay process being overlooked). Thus the reciprocal of the attrition measures the time to collapse in units of the present time per radian.

What has been said here refers to a geon energized by only a single mode of electromagnetic vibration. When several disturbances are present of different frequencies, they will in general leak out at different rates. Moreover, nonstatic components of the gravitational field will furnish a weak nonlinear coupling between the modes, so that relative amplitudes and frequencies on this account also will change gradually with time. The evolution of the system in time is no longer describable as a simple scale transformation. Now it is appropriate to divide up the output of energy, or mass, dM, in the elementary time interval dt, into elementary parts, δdM, each of which refers to that part of the loss which occurs in a specific interval of circular frequency, $\delta\omega$:

$$(\delta dM)/M = -\beta\delta\omega dt. \tag{18}$$

Here β is again a dimensionless quantity, the "attritivity" (attrition constant) of the geon. The attritivity is of course a function of the frequency in question; or more conveniently, it depends on the dimensionless measure of frequency, $\omega^* = MG\omega/c^3$. The fact that a general geon changes the form of its spectral output as time advances prevents us in no way from considering two geons of quite distinct sizes, one of which can be obtained from the other by a simple scale transformation. Both geons then have the same dimensionless attritivity function, $\beta = \beta(\omega^*)$.

The frequency of the outgoing radiation, even in the case of a simple unimodal geon, can only be defined within a latitude of the order $\Delta\omega \sim \alpha\omega$. An analysis of the external electromagnetic field in time, and an analysis in terms of frequency, stand to each other in principle in a mutually exclusive relation. It is only the smallness of α that allows us to speak approximately of the frequency as a function of time. This we do when we go beyond simple transliteration of the function $\mathcal{E}$ = constant $\cos[-\alpha^{-1}\ln(-t)]$ to the form $\mathcal{E}$ = constant $\cos\int\omega dt$, and give $\omega(t) = -1/\alpha t$ the name of "frequency." On this account it is really optional whether the emission spectrum is regarded to be a sharp line of slowly changing frequency, or to be a continuous spectrum concentrated within a region of the order α about a center of gravity of gradually shifting location. In this sense we have a choice whether to describe the decay of a unimodal geon by way of a single number, the attrition, α, or by way of an attritivity function, $\beta(\omega^*)$:

$$\alpha = (\omega^*)^{-1} \int_{\substack{\text{neighborhood} \\ \text{of } \omega^*}} \beta(\omega^*) d\omega^*. \tag{19}$$

In the case of a geon energized by several modes of electromagnetic vibration, the frequencies and amplitudes of which are continually changing by reason of their mutual interaction, the relevant quantity for the description of the slow leakage of energy out of the system is the attritivity.

When a simple monochromatic toroidal geon leaks an amount of radiation energy dE, then its reduced action, $\mathcal{J} = \text{action}/2\pi$, decreases by an amount that satisfies the general relation

$$c^2 dM = dE = \omega d\mathcal{J}. \tag{20}$$

When the decrease becomes substantial in comparison with the original values, then the frequency ω changes substantially. On this account we cannot apply to a geon the special relation, $E=\mathcal{J}\omega$, valid for a harmonic oscillator. Instead, we can note that the circular frequency of the electromagnetic disturbance rises in inverse proportion, $\omega\sim c/\lambda\sim c/(R/a)$, as the mass and radius of the system diminish:

$$GM/c^2\sim R\sim ac/\omega. \tag{21}$$

Multiplying together from (20) and (21) the left- and right-hand sides, respectively, we find

$$Gd(M^2)\sim acd\mathcal{J},$$

or

$$M\sim a^{\frac{1}{2}}(c\mathcal{J}/G)^{\frac{1}{2}}. \tag{22}$$

Thus it is not the mass values themselves, but the squares of the mass values, that are separated in proportion to the intervals of action.

From the mass value (22) follow at once the Eqs. (5) of the introduction for the other physical quantities of the geon as a function of action.

3. THE IDEALIZED SPHERICAL GEON

The simple toroidal geon forms the most elementary object of geon theory much as a simple circular orbit constitutes the first concept of planetary theory. But the simplest physics does not go in the geon case with the simplest mathematics. Toroidal geometry and general relativistic field equations each have their complications, and the mixture requires some time for its analysis.[9] On this account it is natural to look for a geon with spherical symmetry: a rotation-invariant gravitational field, and a spherically symmetric distribution of gravitational source strength, or of stress-energy. Temporarily to adopt a photon point of view, we recognize that each photon orbit is a great circle. Therefor spherical symmetry in the density and flux of photons requires spherical symmetry in the distribution of their angular momentum vectors.

The different elementary disturbances must have different frequencies. If all had the same frequency, they would add coherently to form a single mode of distribution of electromagnetic field strength. But there is no such thing as a nonzero source-free spherically symmetrical electromagnetic field disturbance. Incoherence is essential for sphericity. Let the distribution of field strength be symbolized by the expression

$$\mathcal{E}=\sum E_i(x)\,\sin(\omega_i t+\delta_i),$$

and the distribution of stress-energy by

$$\mathcal{E}^2=\sum_{i,k}\mathcal{E}_i(x)\mathcal{E}_k(x)\,\sin(\omega_i t+\delta_i)\,\sin(\omega_k t+\delta_k).$$

Then distinction of frequencies and randomness of phases is essential to justify the approximation

$$\mathcal{E}^2=\tfrac{1}{2}\textstyle\sum_i \mathcal{E}_i^2(x).$$

The distribution of stress-energy being thus approximately static, it can also be made spherically symmetrical by properly coordinated choices of the elementary solutions $\mathcal{E}_i(x)$.

A geon spherically symmetric in the sense just described is in principle unstable with respect to transformation into a toroidal geon. Tolman showed long ago[8] that two nearly parallel pencils of light attract gravitationally with twice the strength one might have thought when their propagation vectors are oppositely directed, and when similarly directed attract not at all. Consequently a system of randomly oriented circular rays of light will drop to a state of greater stability when half of the angular momentum vectors orient themselves parallel, half antiparallel, to a certain direction in space. Then the number of attractive bonds between orbits will be maximized, and the nullity of the angular momentum will be conserved. Simultaneously there will occur a readjustment in the orbital radii.

The spherical geon, though thus unstable, is in unstable *equilibrium*. We can compare it to a pendulum standing the wrong way up. To envisage such a situation would not be of much use if one had it in mind to discuss the oscillations of the pendulum about its point of support, but is quite acceptable if our aim is to discuss the rigidity of the pendulum rod. Likewise the periods of the various electromagnetic modes of vibration of the geon must be judged very short in comparison with the time of turnover of the system into a toroidal system. In this sense we can talk of the properties of the spherical geon in a reasonably well defined way. It is not necessary to treat all questions of geon stability in order to undertake the problem of the structure of a spherical geon.

The assumed sphericity of the system might appear to be a self-contradictory notion. Spherical symmetry of the static gravitational field, or of the effective refractive index, implies degeneracy of the modes of electromagnetic field oscillation. However, identity of frequency of the various modes, $\mathcal{E}_i(x)$, will rule out the incoherence of the various disturbances so necessary for a spherically uniform mass distribution. Two factors allow us to avoid this degeneracy: slight differences in the scale of the excited modes, and slight departures from spherical symmetry. We consider a disturbance with no radial nodes, and characterize it by the order, l, of the relevant spherical harmonic, and by the azimuthal index number a, of this spherical harmonic. The mode of azimuthal index $a=l$ is concentrated in a toroidal region of major radius $R\sim l\lambda$ and minor radius $\Delta R\sim l^{\frac{1}{2}}\lambda$. The angular dependence of the intensity, for example, shows itself simply in the expression

$$\text{constant}\,|P_l^{(l)}(\theta)e^{il\varphi}|^2=\sin^{2l}\theta$$
$$=(1-\cos^2\theta)^l\doteqdot[1-(\Delta\theta)^2]^l\doteqdot\exp[-(l^{\frac{1}{2}}\Delta\theta)^2],$$

which falls to $\sim(1/e)$th of its value in a distance

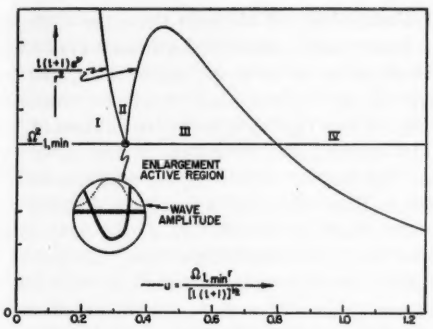


FIG. 2. Case of simple idealized spherical geon; distance dependence of factors in the differential equation for the electromagnetic vector potential,

$$d^2R/dr^{*2}+[\Omega^2-l(l+1)e^\nu/r^2]R=0.$$

The proper solution, R, of lowest circular frequency, Ω/c, for a given high order of spherical harmonic, l, has the qualitative behavior indicated by the dotted line in the diagram. This solution rises exponentially with r in region I, has a single maximum in region II, falls off exponentially in region III, and in region IV oscillates with a very small amplitude (leakage wave). There exist also solutions of the same order, l, but of higher frequency, Ω, corresponding to a higher position for the horizontal line in the figure. The square of the wave number is represented by the distance measured positively upwards from the curve to the horizontal line. An increase of $l^*=[l(l+1)]^{\frac{1}{2}}$ and of Ω by the same factor leaves all turning points unchanged. However, the height of the barrier against leakage is increased as l^{*2}, so that the exponent of the penetration factor is proportional to l^*. The size of the active region is very small on the scale of the diagram, being proportional only to $l^{*\frac{1}{2}}$; hence the enlargement of this region, presented more in detail in Fig. 3. In the present figure the left-hand section of the curve up to $u=b\doteq\frac{1}{2}$ is given by the expression b^2/u^2; and the right-hand region by $(1/u^2)[1-b(1-b^2/u)]$.

$\Delta R=R\Delta\theta\sim R/l^{\frac{1}{2}}$, measured angularwise. Radialwise let us consider a fixed and pre-existing metric (Fig. 2). Then for each angular order of l there exists a minimum circular frequency, $\omega=c\Omega_{l,\min}$, such that the corresponding mode shows only a single radial maximum in the active region. For different l's the maxima overlap one another. Relative to the coordinate, $r=R$, of this maximum, the quantity $1/\Omega\equiv\lambda$ has a value of the order of R/l. The width of the region of maximum activity is of the order of $R^{\frac{1}{2}}\lambda^{\frac{1}{2}}\sim l^{\frac{1}{2}}\lambda\sim R/l^{\frac{1}{2}}$, as follows by simple JWKB analysis from the circumstance that the square of the wave number changes linearly with departure from the point of maximum wave number with a slope, $d(\text{wave number})^2/dr$, of the order of $\Omega^2/R\sim l^2/R^3$. Thus a sequence of l values yields a family of modes, of which the stress-energy is all concentrated in the same region, $r=R$. To have the spatial spread of energy also nearly the same for all modes, we have only to ask that the range, Δl, of the orders of the activated modes should not exceed some reasonable fraction of l itself. As the other factor freeing the oscillations from degeneracy we must admit some departure of the average gravitational field from exact spherical symmetry. A slight ellipticity will split up into $l+1$ separate frequencies the otherwise $(2l+1)$-fold degenerate vibrations of order l. Counting together both possibilities at our disposal, we have of the order of l^2 modes with distinct frequencies. On the other hand, to obtain a distribution of energy that is completely symmetric in angle it is sufficient to add the squares of the $(2l+1)$ normalized spherical harmonics of order l, according to the well-known completeness theorem of spherical harmonics. Moreover, we do not demand for the geon a completely smooth distribution of the energy over the surface, for the long-range character of the law of gravitation smooths out in the field the minor irregularities that were present in the mass density. Consequently we can actually do with disturbances having considerably fewer than $2l+1$ values of the azimuthal index, a. Comparing $2l+1$ with $\sim l^2$, we conclude that we have more than adequate margin to obtain the postulated uniformity of energy distribution with nondegenerate modes.

For the purposes of calculations it is justifiable to idealize the picture we have just formed. The radial distribution of energy does not differ importantly from one of the excited modes to another; consequently we treat the radial factors in all the distributions as identical to that calculated for a single fixed value of l. The total angular distribution is nearly uniform, so we take it to be exactly uniform; i.e., we take the oscillations of the $2l+1$ values of a belonging to the given order, l, as normalized to identical energy contents. The only consequence of exact spherical symmetry that we do not accept is *coherence* of the $2l+1$ modes. In the actual "spherical" geon, we know that the frequencies are distinct; so in the idealized spherical geon, where legalistically the frequencies come out identical, we disregard this circumstance, and add energy densities according to the law for incoherent disturbances.

To proceed to details, we make Schwarzschild's choice of polar coordinate system, r, θ, φ, such that elementary distances at right angles to the radius vector are represented by the usual geometrical formula. Then dr gives proper distances, ds, in the radial direction only after correction by a certain r-dependent factor. Likewise dt, or rather the cotime, $dT=cdt$, gives correctly intervals of proper cotime, $d\tau$, only after multiplication by another r-dependent factor. Both correction factors are independent of T because the gravitational field is static. One factor is the reciprocal of the other in the well-known Schwarzschild solution for the metric about a point mass, but no such simple relation holds in the interior of an object with the distributed mass of a geon:

$$(ds)^2=-(d\tau)^2=g_{\alpha\beta}dx^\alpha dx^\beta=e^{\lambda(r)}(dr)^2 +r^2[(d\theta)^2+(\sin\theta d\varphi)^2]-e^{\nu(r)}(dT)^2. \quad (23)$$

As solution of the electromagnetic field equations it is reasonable to look for a periodic function of time, multiplied by a function of r, multiplied by a function of angle generated from the spherical harmonic $(\sin\theta)^l\exp(il\varphi)$. Such a disturbance is easily visualized in terms of a toroidal concentration of electromagnetic

energy. The other elementary disturbances that we need to give a spherically symmetric distribution of energy are readily generated from this mode of vibration by the operations of the rotation group. This circumstance means that we can consider as starting point one mode as well as another. Mathematically simplest is not the choice just named, with $a=l$, but the spherical harmonic azimuthal index $a=0$, for in this case the dependence upon φ disappears. Thus we look for a solution of the form

$$A_r=A_\theta=A_T=0;$$
$$A_\varphi=(\sin\Omega T)R(r)\sin\theta(d/d\theta)P_l(\cos\theta), \quad (24)$$

where either group theory or the relevant differential equation tells us at once the form of the angular factor. The electric field points always in the φ direction and the magnetic field and the Poynting vector lie always in the meridian plane:

$$E_\varphi\sim F^{T\varphi}=-e^{-\nu}(r\sin\theta)^{-2}\partial A_\varphi/\partial T,$$
$$-H_\theta\sim F^{r\varphi}=e^{-\lambda}(r\sin\theta)^{-2}\partial A_\varphi/\partial r,$$
$$H_r\sim F^{\theta\varphi}=(r^2\sin\theta)^{-2}\partial A_\varphi/\partial\theta. \quad (25)$$

In principle we should consider modes which differ from (25) by the interchange of the role of electric and magnetic vectors. To do so would make no difference in the final averaged energy and stress densities, for divergence-free relativity is, despite its appearance, completely symmetric between electricity and magnetism. Consequently we do not need to sum over polarizations. We deal with a standing wave with flow of energy back and forth from north to south pole of the sphere, concentrated in the present case in an active region between R and $R+\Delta R$. We have the superposition of a number of toroidal disturbances, each passing through the north and south poles, with all azimuths of orientation of the tori weighted equally. In line with this picture, A_φ is stronger in the region of overlap near the poles than it is at the equator.

Of the electromagnetic field equation (3), three components unite in saying that A_φ should be independent of φ, and in saying only this: while the remaining, or φ, component, makes the statement

$$(\partial/\partial r)\exp[(\nu-\lambda)/2](\sin\theta)^{-1}\partial A_\varphi/\partial r$$
$$+(\partial/\partial\theta)\exp[(\nu+\lambda)/2](r^2\sin\theta)^{-1}\partial A_\varphi/\partial\theta$$
$$-\exp[-(\nu-\lambda)/2](\sin\theta)^{-1}\partial^2 A_\varphi/\partial T^2=0,$$

whence the radial wave equation

$$d^2R/dr^{*2}+[\Omega^2-l(l+1)e^\nu/r^2]R=0. \quad (26)$$

Here dr^* is an abbreviation for

$$dr^*=[\exp(\lambda-\nu)/2]dr.$$

Let us imagine that we have calculated the time-average value, $\langle T_i{}^k{}_{(I)}\rangle$, of the electromagnetic stress-momentum-energy tensor of the disturbance under discussion: mode number I. This tensor depends only on r and θ. Let there be altogether N modes of different frequencies excited, all of about the same strength and properties as mode I, but differing from it in orientation. Let the distribution of stress, momentum, and energy in a typical one of these disturbances be visualized by mapping out on a rigid sphere the distribution of mode I, and then rotating the north pole of this sphere to an angle α with respect to its original position, and to an azimuth β. Thus $\langle T_i{}^k{}_{(D)}\rangle$, for a typical disturbance, D, is a tensor function of r, θ, and φ, the functional dependence being completely known in principle as soon as we know the angles α and β associated with the mode D. Let us sum the values at a typical point, θ, φ, of the energy density over all N modes, obtaining

$$\langle T_T{}^T\rangle=(N/4\pi)\int\int\langle T_T{}^T{}_{(D)}\rangle\sin\alpha d\alpha d\beta. \quad (27)$$

This result cannot depend upon θ and φ, because of the spherical symmetry. Consequently the point of evaluation can be taken to lie at the original location of the north pole, $\theta=0$. Here the energy density of disturbance D is the same as was the energy density of disturbance I at the polar angle $\theta=\alpha$. Consequently (27) reduces to the form

$$\langle T_T{}^T\rangle=(N/2)\int\langle T_T{}^T{}_{(I)}\rangle\sin\theta d\theta. \quad (28)$$

Similarly we obtain for the radial tension

$$\langle T_r{}^r\rangle=(N/2)\int\langle T_r{}^r{}_{(I)}\rangle\sin\theta d\theta. \quad (29)$$

In evaluating the tangential components of the tension we have to proceed more carefully, for the definition of the relevant directions differs between the original mode and the rotated modes. Consequently we have to employ the standard rules for transformation of the components of a tensor. Fortunately all the transformations are carried out in the local tangent space, and the rules for Cartesian tensors apply:

$$\langle T_\theta{}^\theta\rangle=(N/4\pi)\int[\langle T_\theta{}^\theta\rangle\cos\beta\cos\beta$$
$$+\langle T_\theta{}^\varphi{}_{(I)}+T_\varphi{}^\theta{}_{(I)}\rangle\sin\beta\cos\beta$$
$$+\langle T_\varphi{}^\varphi{}_{(I)}\rangle\sin\beta\sin\beta]\sin\alpha d\alpha d\beta$$
$$=(N/4)\int\langle T_\theta{}^\theta{}_{(I)}+T_\varphi{}^\varphi{}_{(I)}\rangle\sin\theta d\theta. \quad (30)$$

The expression for $\langle T_\varphi{}^\varphi\rangle$ is identical to this. All other components of the total stress-energy tensor have mixed indices and must vanish on account of one or another symmetry argument. Thus $\langle T_\theta{}^T\rangle$ and $\langle T_\varphi{}^T\rangle$ describe the tangential flow of energy, which has to vanish; the shears $\langle T_r{}^\theta\rangle$, $\langle T_r{}^\varphi\rangle$, $\langle T_\theta{}^\varphi\rangle$ must be null; and the component $\langle T_r{}^T\rangle$ represents the net radial flow of

energy, which in a static situation must also be zero. Whether the geon leaks no energy at all, or only a very little energy, depends upon whether we impose upon the electromagnetic field at large distances the requirement that it be a standing wave, or a pure outgoing wave. Which choice we make has negligible effect upon the structure of the geon. We already made our choice when we took the circular frequency, $c\Omega$, to be real and assumed the product representation (24) of the vector potential. From the radial equation (26) it then follows automatically that $R(r)$ must represent a standing wave, and that the radial flux must vanish. In summary, we have for the nonzero components of the average value of the total electromagnetic stress-energy tensor (2) the following values:

$$\langle T_r{}^r\rangle = (N/2)\int \sin\theta d\theta (8\pi)^{-1}(\{r\varphi\}-\{T\varphi\}-\{\theta\varphi\})$$

$$=[r\varphi]-[T\varphi]-[\theta\varphi],$$

$$\langle T_\theta{}^\theta\rangle = \langle T_\varphi{}^\varphi\rangle = (N/2)\int \sin\theta d\theta (8\pi)^{-1}\{\theta\varphi\}$$

$$=[\theta\varphi],$$

$$\langle T_T{}^T\rangle = (N/2)\int \sin\theta d\theta (8\pi)^{-1}(\{T\varphi\}-\{r\varphi\}-\{\theta\varphi\})$$

$$=[T\varphi]-[r\varphi]-[\theta\varphi]. \tag{31a}$$

Here we used as abbreviations:

$$\{r\varphi\}\equiv\langle F_{r\varphi}F^{r\varphi}\rangle = e^{-\lambda}(r\sin\theta)^{-2}\langle(\partial A_\varphi/\partial r)^2\rangle\sim\langle H_\theta{}^2\rangle;$$

$$\{\theta\varphi\}\equiv\langle F_{\theta\varphi}F^{\theta\varphi}\rangle = (r^2\sin\theta)^{-2}\langle(\partial A_\varphi/\partial\theta)^2\rangle\sim\langle H_r{}^2\rangle;$$

$$\{T\varphi\}\equiv\langle F_{T\varphi}F^{T\varphi}\rangle = -e^{-\nu}(r\sin\theta)^{-2}$$

$$\times\langle(\partial A_\varphi/\partial T)^2\rangle\sim-\langle E_\varphi{}^2\rangle,$$

and

$$[\ \ ]=(N/16\pi)\int\{\ \ \}\sin\theta d\theta;\ \text{thus}$$

$$[r\varphi]=(N/16\pi)l(l+1)(2l+1)^{-1}e^{-\lambda}(dR/rdr)^2;$$

$$[\theta\varphi]=(N/16\pi)l^2(l+1)^2(2l+1)^{-1}(R/r^2)^2;$$

$$[T\varphi]=-(N/16\pi)l(l+1)(2l+1)^{-1}e^{-\nu}(\Omega R/r)^2. \tag{31b}$$

Having an equation for the influence of gravitation upon the electromagnetic disturbance (26), and expressions (31) for the stress-energy density of this disturbance, we can now complete the circle of the self-consistent system and write down the equations

$$G_{ik}=(8\pi G/c^4)T_{ik} \tag{32}$$

for the determination of the gravitational potentials. In the mixed covariant-contravariant notation (which gives $T_4{}^4$ a sign opposite to T_{44}), we have[11]

$$G_r{}^r = e^{-\lambda}(r^{-1}d\nu/dr+r^{-2})-r^{-2},$$

$$G_\theta{}^\theta = G_\varphi{}^\varphi = \tfrac{1}{2}e^{-\lambda}[d^2\nu/dr^2+\tfrac{1}{2}(d\nu/dr)^2+r^{-1}d\nu/dr$$

$$-r^{-1}d\lambda/dr-\tfrac{1}{2}(d\nu/dr)(d\lambda/dr)],$$

$$G_T{}^T = e^{-\lambda}(r^{-2}-r^{-1}d\lambda/dr)-r^{-2}. \tag{33}$$

The field equation $G_\theta{}^\theta=(8\pi G/c^4)T_\theta{}^\theta$ is identical with the equation for $G_\varphi{}^\varphi$ and can be disregarded, for it says nothing in addition to the equations for $G_T{}^T$ and $G_r{}^r$. One has only to differentiate with respect to r the equation for $G_r{}^r$, add to it a proper linear combination of the equations for $G_r{}^r$ and $G_T{}^T$, and employ the wave equation (26) to get the content of the equation for $G_\theta{}^\theta$. This result only says that the tangential tensions in a spherical shell have to be balanced by the radial pressure gradient. We therefore end up with two equations of the first order for the two unknown functions λ (= twice the number of napiers of change of scale in radial distances) and ν (= twice the number of napiers of change of scale in time measurements):

$$e^{-\lambda}(r^{-1}d\nu/dr+r^{-2})-r^{-2}=[l(l+1)GN/2(2l+1)c^4]$$

$$\times[e^{-\lambda}(dR/rdr)^2+e^{-\nu}(\Omega R/r)^2-l(l+1)(R/r^2)^2]; \tag{34}$$

$$e^{-\lambda}(r^{-2}-r^{-1}d\lambda/dr)-r^{-2}=[l(l+1)GN/2(2l+1)c^4]$$

$$\times[-e^{-\lambda}(dR/rdr)^2-e^{-\nu}(\Omega R/r)^2-l(l+1)(R/r^2)^2]. \tag{35}$$

The scale invariance of classical geon theory shows itself at once. For geons of the same index l, but different sizes and masses and therefore different circular frequencies $c\Omega$, we introduce the same dimensionless measure of radial coordinate, $\rho=\Omega r$. Also we define the dimensionless measure of potential

$$f(\rho)=[l(l+1)GN/2(2l+1)c^4]^{\frac{1}{2}}\Omega R(r). \tag{36}$$

Furthermore we recall that the Schwarzschild solution for the field of a point mass has the form $e^{-\lambda}=e^{\nu}=1-2r^{-1}L_0$, where $L_0=GM/c^2$ is a measure of the mass of the object. We have to expect a similar result for the metric around the geon at distances where the electromagnetic field has fallen exponentially to negligible values, but where the gravitational field may be still quite strong. Consequently we shall write

$$e^{-\lambda}\equiv 1-2\rho^{-1}L(\rho),$$

$$e^{\lambda+\nu}\equiv Q^2(\rho),$$

$$e^{\nu}=[1-2\rho^{-1}L(\rho)]Q^2(\rho). \tag{37}$$

Here the dimensionless measure, $L(\rho)$, of mass inside the radius r is nearly zero from $\rho=0$ to a value of ρ close to the inner surface of the active region of the geon, a value of the order of

$$\rho\sim l^*\equiv[l(l+1)]^{\frac{1}{2}};$$

then $L(\rho)$ rises quickly over a range of ρ of the order

[11] See for example L. Landau and E. Lifschitz, *The Classical Theory of Fields*, translated by M. Hamermesh (Addison-Wesley Press, Cambridge, 1951).

$\Delta\rho \sim l^{*\frac{1}{2}}$; and then $L(\rho)$ stays essentially constant from this point to very large distances. The correction factor, Q, is likewise essentially constant inside and outside the geon, and only makes a sudden rise in the active region. In these notations the geon is described by the self-consistent solution of three equations: the wave equation

$$d^2f/d\rho^{*2}+[1-(l^*Q/\rho)^2(1-2L/\rho)]f=0, \quad (38)$$

where $d\rho^*$ is an abbreviation for

$$d\rho^*=Q^{-1}(1-2L/\rho)^{-1}d\rho; \quad (39)$$

and the two field equations, of which the first gives the change of mass with distance:

$$dL/d\rho^*=(1/2Q)[f^2+(df/d\rho^*)^2 + (l^*Qf/\rho)^2(1-2L/\rho)], \quad (40)$$

and the second gives the variation of the factor Q:

$$dQ/d\rho^*=(\rho-2L)^{-1}[f^2+(df/d\rho^*)^2]. \quad (41)$$

It is an interesting feature of the system of equations (38), (40), (41) that they still permit the possibility of a change of scale of distance without a change of form:

$$\begin{aligned} \rho &= b\rho_1, \\ Q &= bQ_1, \\ \rho^* &= \rho_1^*, \\ L &= bL_1, \\ f &= b^{\frac{1}{2}}f_1. \end{aligned} \quad (42)$$

Now $g_{rr}=e^{\lambda}$ is already so normalized, according to Eq. (37), that it goes over to the Euclidean value of 1 at very large distances. Consequently we wish the corresponding cotime factor, e^{ν}, also to go to unity. Another value is perfectly possible, but it would correspond to an unhappy choice for the value of the speed of light. Thus, according to (37), we want the factor Q to go to unity at large distances. This normalization means, according to the law (41) of increase of Q, that Q must start off at the origin with a value less than 1. But how much less we do not know until we have integrated the differential equation. Accordingly, we distinguish between the solution Q, L, f that we want and the solution Q_1, L_1, f_1 that we get from our numerical integration of the equations of the self-consistent field. In both solutions the mass factor, L or L_1, and the field factor, f or f_1, start at zero and are well behaved at large distances. In the former Q starts at a value less than one and rises to unity; in the latter we start, for convenience, at $Q_1=1$ and arrive at large distances at a value greater than one. This value then defines the scale factor required in (42) to go from the preliminary solution to the desired solution:

$$b=1/Q_1(\infty); \quad Q_1(0)=1. \quad (43)$$

In preparing for the numerical integration of the geon equations we note that the behavior of the solutions near the origin is obvious and not very interesting

$$\begin{aligned} f_1 &= F\rho_1^{l+1}+\cdots, \\ L_1 &= 2^{-1}(l+1)F^2\rho_1^{2l+1}+\cdots, \\ Q_1 &= 1+(2l)^{-1}(l+1)^2F^2\rho_1^{2l}+\cdots. \end{aligned} \quad (44)$$

What is important is the circumstance that there is no adjustable constant except the starting value of the field strength. This field strength, or the constant F, must be so chosen as to give the geon stability. We have to do with an eigenvalue, but an eigenvalue of a nonlinear system of equations.

The key to the eigenvalue problem is the wave equation (38). Where the expression in square brackets is positive, there the solution is oscillatory; where it is negative, there the solution rises or falls approximately exponentially. This expression has the following behavior when the factor F in the field strength has been chosen to give the geon stability. (1) At distances less than those that characterize the active region of the geon, $L_1 \doteq 0$, $Q_1 \doteq 1$, and the bracket has the approximate form $[1-(l^*/\rho_1)^2]$. Thus, in moving out from $\rho_1=0$ to ρ_1 near l^*, we have until the very end a strongly negative bracket, and therefore a solution that rises approximately exponentially. Consequently we have yet to arrive at the zone of maximum activity. (2) In the active zone near $\rho_1=l^*$ there occurs a quick rise in Q_1 and L_1, with very little change in the radial coordinate. When the field strength is properly chosen, these changes that follow from (40) and (41) are such that the square bracket is positive over a limited range of ρ. Thus f rises to a maximum and starts to fall off exponentially as ρ_1 passes out of the region of oscillation, provided that the field strength is great enough to give the geon the requisite mass for stability in a single wave zone. (3) For values of ρ_1 just slightly larger than l^*, the mass factor L_1 and the quantity Q_1 have attained essentially constant values. These values are such that $(1-2L_1/l^*)$ is much less than unity, but Q_1^2 is much greater than unity, and the product of these two factors also exceeds unity by a considerable margin. Moreover, as ρ_1 increases, the factor $(1-2L_1/\rho_1)$ evidently rises rapidly percentagewise at first, then levels off to a saturation value. The factor $(Q_1l^*/\rho_1)^2$, on the other hand, falls off in a more nearly uniform way. Consequently the product of these two factors rises as we leave the active zone behind, reaches a maximum when the increase, ρ_1-l^*, of the radial coordinate is some substantial fraction of l^*, then falls off and passes through the value unity for ρ_1-l^* of the order of l^* (Fig. 2). The product of factors under discussion in the present classical problem evidently plays the part of the potential in a typical problem of quantum mechanics. The analogy with the theory of alpha decay is complete even to topological identity of the two forms of barrier to be penetrated. In the barrier in the present problem the field factor f of an eigenfunction falls off exponentially by a number of napiers of the order of l^*.

(4) With further increase of the radial coordinate beyond the point of emergence from the barrier, oscillation sets in. The amplitude is of course exceedingly small compared to the amplitude inside the geon. The wavelength is at first long but eventually settles down to the value appropriate to a disturbance of the given frequency in free space:

$$f/\text{constant} \doteq \sin(\rho^*+\delta) = \sin(Q_1(\infty)\rho_1+\delta) = \sin(\rho+\delta) = \sin(\Omega r+\delta). \quad (45)$$

To go further, we seem to have to obtain for each conceivable value of the index number l a self consistent numerical solution of three nonlinear equations. However, there exists in addition to the exact scaling law already exploited another and distinct scaling law, valid asymptotically for large l^*, which reduces all these problems to a single one. Similar scaling laws are familiar from other parts of mathematical physics. In quantum mechanics one reduces all hydrogenic atoms to one problem by appropriate choice of scale of length. However, then the radial wave function for the lowest state for each value of l has a different mathematical expression, $\{[(l+1)/2]^{2l+3}(2l+2)!\}^{-\frac{1}{2}}\rho^l\exp[-\rho/(l+1)]$. Yet simple analysis of the behavior of this function in the neighborhood of its maximum, $\rho=l(l+1)$, shows that it has a form,

$$\simeq[\pi l(l+1)^2]^{-\frac{1}{4}}\exp[-(\Delta\rho)^2/2l(l+1)^2],$$

which is the same up to a scale change for all values of l. A similar scaling principle applies to the geon, with two differences: the reduced field factor, f, does not have the form of a harmonic oscillator wave function; and the relevant length variable is not $x=\Delta\rho/l^{\frac{1}{3}}(l+1)$, as in the hydrogenic problem, but

$$x=(\rho^*-l^*)l^{*-\frac{1}{3}}. \quad (46)$$

The whole of the active region of the geon will be described by a range of x of the order of unity.

To bring this similarity transformation into evidence, we consider large values of l^*, and expand the relevant quantities in inverse powers of $l^{*\frac{1}{3}}$:

$$\begin{aligned} d\rho^* &\equiv l^{*\frac{1}{3}}dx, \\ \rho_1 &= l^*+l^{*\frac{1}{3}}r_0(x)+\cdots, \\ L_1 &= l^*\lambda_0(x)+l^{*\frac{1}{3}}\lambda_1(x)+l^{*\frac{2}{3}}\lambda_2(x)+\cdots, \\ Q_1 &= [1/k(x)]+l^{*-\frac{1}{3}}q_1(x)+l^{*-\frac{2}{3}}q_2(x)+\cdots, \\ f_1 &= l^{*\frac{1}{3}}\varphi(x)+\varphi_1(x)+l^{*-\frac{1}{3}}\varphi_2(x)+\cdots, \\ [1-(Q_1l^*/\rho_1)^2(1-2L_1/\rho_1)] &= l^{*-\frac{1}{3}}j(x)k(x)+\cdots. \end{aligned} \quad (47)$$

Inserting these expressions into the system of equations (38), (40), (41), and identifying coefficients of like powers of l^*, we find from the lowest relevant terms the wave equation

$$d^2\varphi/dx^2+j(x)k(x)\varphi(x)=0, \quad (48)$$

and the two field equations:

$$dk/dx+\varphi^2=0, \quad (49)$$

$$dj/dx=3-[1+(d\varphi/dx)^2]/k^2. \quad (50)$$

If we can find by integration of these three equations the reduced field factor $\varphi(x)$ and the supplementary time scale correction factor $k(x)$ and the reduced oscillation factor $j(x)k(x)$, then we can determine the other leading terms in expressions (47) by simple calculation:

$$dr_0/dx=k(x),$$

$$\lambda_0=(1-k^2)/2. \quad (51)$$

We seek a solution of (48), (49), and (50) with the following properties. The field factor $\varphi(x)$ tends to zero both for large positive x and for large negative x. The contraction factor $k(x)$ is unity for large negative x, and falls in the active region near $x=0$, and for large positive x approaches a value which is less than $(1/3)^{\frac{1}{2}}=0.577$ but still positive. The quantity $j(x)$ is very large and negative for large negative x. It rises with increasing x with a positive slope of two until x reaches the vicinity of the active region. There j succeeds in becoming positive for a limited range of x in the neighborhood of $x=0$. For larger x, the slope dj/dx approaches the negative value $3-k^{-2}(\infty)$. Thus j falls off again to $-\infty$. The oscillation factor, the product $j(x)k(x)$, is positive in only a limited range of x.

There exist a number of solutions having the desired behavior, distinguished from one another by an integer, s, which represents the sum of the number of maxima and minima in the field factor, $\varphi(x)$, in the active region. This integer has the value one for the simplest type of idealized spherical geon. Higher values of s also represent physically acceptable geons. We make the choice between one solution and another when in the integration of the equations we pick one starting magnitude of $\varphi(x)$ or another. The solution $s=1$ is characterized by the largest field strength, for in one-half wave enough energy and mass has to accumulate to hold the geon together. A comparable amount of mass belongs to a geon of higher s, but the energy is distributed over a larger number of half waves and on this account the concentration of energy and the field strength are weaker.

We can summarize the eigenvalue characteristics of our system of equations as follows. We accept that for large negative x, the factors j and k have the form $k(x)\doteq 1$ and $j(x)\doteq 2x$, and that $\varphi(x)$ is approximately the exponentially rising solution of

$$d^2\varphi/dx^2=2x\varphi; \quad (52)$$

thus,

$$\varphi\doteq A(-2x)^{-\frac{1}{4}}\exp[-(-2x)^{\frac{3}{2}}/3]. \quad (53)$$

Here the constant A, the "amplitude factor," is the sole quantity at our disposal in selecting the character of the solution of (48), (49), (50). If A is chosen very small in comparison with unity, then (49) says that $k(x)$ stays close to 1 for a long range of x, and (50) says that $j(x)$ continues to behave nearly as $2x$ for a considerable range of positive values of x. There likewise (52) remains a good approximation, and the solution, essentially

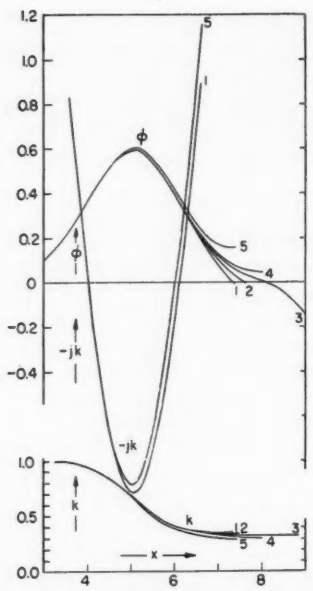


FIG. 3. Results of numerical integration of the differential equations (48), (49), (50) for the simple idealized self-consistent spherical geon, giving a more detailed account of the active region shown in the circle in Fig. 2. Here φ is the dimensionless measure of the radial factor in the electromagnetic potential, jk is the oscillation factor in the differential equation $d^2\varphi/dx^2+jk\varphi=0$, and k is a metric correction factor. The range of integration from start at $x=-4$ up to $x=3$ is not shown owing to the inappreciable difference there between the curves with nearly identical starting values: $\varphi(-4)\equiv\varphi_0$; $\varphi_0\times10^4=1.025$ (case 1); 1.028125 (2); 1.03 (3); 1.03125 (4); 1.0375 (5). The eigenvalue evidently lies between cases (3) and (4). For any given geon mass in the classical range, the curves permit one in principle to find field strengths, metric, and all other relevant physical quantities in the active region of the geon.

an Airy function, goes over from the exponential character (53) to the oscillatory form

$$\varphi\doteq 2A\,(2x)^{-\frac{1}{4}}\sin[(2x)^{\frac{3}{2}}/3+(\pi/4)], \tag{54}$$

passing through the value

$$\varphi\doteq 6^{5/6}2^{-1}\pi^{-\frac{1}{2}}\Gamma(4/3)A$$

at $x=0$. Averaging over oscillations in this first part of the active region, we have

$$dk/dx\doteq-\langle\varphi^2\rangle\doteq-\sqrt{2}A^2/x^{\frac{1}{2}},$$

or

$$k\doteq 1-2^{\frac{1}{2}}A^2x^{\frac{1}{2}};$$

and

$$dj/dx\doteq 2-2^{\frac{1}{2}}3A^2x^{\frac{1}{2}},$$

or

$$j\doteq 2x-2^{5/2}A^2x^{\frac{3}{2}}. \tag{55}$$

These approximations begin to fail when x becomes of the order $1/A^4$. Thereafter we quickly come to the end of the active region. (1) If the amplitude factor has the value, A_s, appropriate to the sth proper solution, then this field factor, $\varphi(x)$, falls off exponentially. (2) If the amplitude factor has a value so little greater than A_s that the difference, $A-A_s$, is small compared to the distance, $A_{s-1}-A_s$, to the next eigenvalue, then there is not quite opportunity to finish the last oscillation in the active region. The field factor starts to fall off after what would have been the last maximum but instead of continuing to fall towards zero, reaches a minimum and then commences to rise exponentially (Fig. 3). The contribution to the mass therefore also rises exponentially. The scale factor k of Eq. (49), instead of decreasing to a reasonable value between 0 and 0.577, continues to fall, and goes to zero at a certain critical value, x^*, as (x^*-x) multiplied by a slowly varying function which has qualitatively the character of a power of $\ln 1/(x^*-x)$. The quantity j therefore becomes singular, and goes to negative infinity as $-(x^*-x)^{-3}$ multiplied again by a slowly varying factor of logarithmic character. Finally, the field factor $\varphi(x)$ goes to infinity qualitatively with much the slowly varying behavior of a power of a logarithm of $[1/(x^*-x)]$. Thus the integration stops at x^*. This limit will be the further beyond the active region, the closer the amplitude factor lies to A_s. (3) As the difference $A-A_s$ becomes larger, the singular point of the integration moves inwards towards what one might still appropriately call the active zone. For a certain A the point x^* has come in to its maximum extent. With further increase in A this critical point moves out again. At or near the value of A for which this turnabout occurs, the sign of φ at the singularity also changes; i.e., the singularity swallows up the last node of the wave function. Thus, for A just less than the critical value, $\varphi(x)$ is described by a quasi-logarithmic function that goes to infinity with one sign; for A just above the critical value, by a singular quasi-logarithmic function of the opposite sign; and for A just right to annihilate a node, by (x^*-x) times a quasi-logarithmic function of (x^*-x). In this particular case the scale factor $k(x)$ goes as $(x^*-x)^3$ times a pseudo-logarithmic function, and the oscillation factor jk goes as $-(x^*-x)^{-2}$ times a pseudo-logarithmic function. (4) As A approaches the proper amplitude factor, A_{s-1}, of the next solution, the singularity moves out towards $x^*=+\infty$ and finally disappears. (5) With further increase in A the sequence (1), (2), (3), (4) repeats itself until finally there are no more nodes in the wave function to be swallowed. At this stage we have passed the last eigenvalue, A_1.

It is appropriate to notice that the system of equations (48), (49), (50) does not contain explicitly the argument x. Hence, if $\varphi(x)$, $k(x)$, $j(x)$ is a solution, so is $\varphi(x+a)$, $k(x+a)$, $j(x+a)$. Consequently there is a certain indeterminacy in the start of the integration process. We remove this ambiguity and completely—though indirectly—define the origin of x when we adopt as convention the asymptotic formula $j(x)=-2x$ and the corresponding asymptotic equations (52) and (53).

With this convention for origin of x, the reduced equations of the spherical geon were integrated numerically to determine approximately the first proper func-

tion and first proper value. Thanks are due to Mr. Arthur Komar for checking the algebra and to Mr. Robert Goerss for carrying out the computations on an International Business Machines card-programmed electronic calculator. The logical scheme of the integration is outlined in Fig. 4, where each line symbolizes a successive stage in the calculation, and where Δ represents the size, 0.05, of the interval of x used between one step and the next. The integration was started at $x=-4$ with values derived from the asymptotic formula (53):

$$\varphi_0\equiv\varphi(-4)=\text{arbitrary},$$
$$\delta\varphi_{\frac{1}{2}}=8^{\frac{1}{2}}\varphi_0\Delta,$$
$$k_0=1,$$
$$j_0=-8. \tag{56}$$

Figure 3 presents the results of the integrations. The first eigenvalue was found to lie between those amplitude factors, A, that correspond to initial values, at $x=-4$, between $\varphi_0=1.03000\times10^{-4}$ and $\varphi_0=1.03125\times10^{-4}$. The active region ($jk>0$) did not begin until $x=4.05$, and reached only to $x=6.12$. The field factor, $\varphi(x)$, reached a peak value of 0.59 for x about equal to 5.1. The scale factor, $k(x)$, approached an asymptotic value, $k(\infty)$, of approximately 0.33.

These curves and numbers give us essentially all the information we need to determine the structure of a simple idealized spherical geon with only one maximum, $s=1$, in the electromagnetic potential and with number of nodes over the surface equal to any arbitrary large number l. Thus, having chosen an l, we calculate $l^*=[l(l+1)]^{\frac{1}{2}}$. This quantity represents the approximate radial coordinate, ρ_1, of the active zone. Then, from our curves and Eqs. (47) and (51) we find the relation between the dimensionless measures of distance, ρ_1; of mass out to a given distance, L_1; of metric correction factor, Q_1; and of potential, f_1. In particular we have for the asymptotic value of Q_1 the value $1/b\doteq1/0.33$ and for the asymptotic value of L_1 the value $L_1=l^*(1-b^2)/2\doteq0.45l^*$. Next, we make the scale change of Eqs. (42) and (43) in order to have a metric

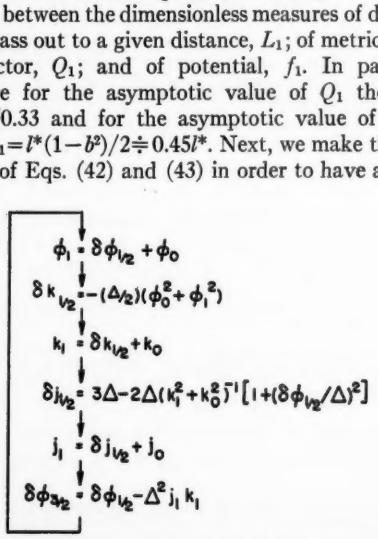

$\phi_1 = \delta\phi_{1/2} + \phi_0$
$\delta k_{1/2} = -(\Delta/2)(\phi_0^2 + \phi_1^2)$
$k_1 = \delta k_{1/2} + k_0$
$\delta j_{1/2} = 3\Delta - 2\Delta(k_1^2 + k_0^2)^{-1}[1 + (\delta\phi_{1/2}/\Delta)^2]$
$j_1 = \delta j_{1/2} + j_0$
$\delta\phi_{3/2} = \delta\phi_{1/2} - \Delta^2 j_1 k_1$

FIG. 4. Flow diagram of numerical integration. In the last term in the last equation there should appear—and did appear in the numerical calculations—a factor ϕ_1.

that approaches the Euclidean values at large distances. The scale factor is $b\doteq0.33$. In the new dimensionless units the radial coordinate of the active zone is about $0.33l^*$. The metric correction factor Q has the value $b\doteq0.33$ in the inner inactive part of the geon, rises in the active zone, and has outside a value 1.00. The corresponding new dimensionless measure of mass, L, is essentially zero in the inner inactive region, and outside the active region has the value $L\doteq(b/2)(1-b^2)l^*=0.15l^*$. Finally, we transform to cgs units of measure via (36) and (37) and the related discussion. A clock ticks at the center of the geon at only about 33 percent of the rate of an identical clock far away from the geon. In terms of the frequency, Ωc, of electromagnetic radiation observed to come from the geon, the radius of the active zone of the system is described by a coordinate

$$r=R\doteq0.33l^*/\Omega. \tag{57}$$

The mass is

$$M\doteq0.15c^2l^*/G\Omega. \tag{58}$$

Let an observer far from the geon have a clock that flashes every second, and let an identical clock be placed at the center of the geon. The flashes of light from this second clock will reach the observer, not every second, but about every $b^{-1}\doteq(0.33)^{-1}=3$ seconds.

The mass, circular frequency, and radius of the spherical geon can be expressed in terms of the reduced action, $\mathcal{J}(=\text{action}/2\pi)$, by way of the relation

$$c^2dM=c\Omega d\mathcal{J}, \tag{59}$$

thus:

$$M\doteq0.54(l^*\mathcal{J}c/G)^{\frac{1}{2}};$$
$$c\Omega\doteq0.27(l^*c^5/G\mathcal{J})^{\frac{1}{2}};$$
$$R\doteq1.2(l^*G\mathcal{J}/c^3)^{\frac{1}{2}}. \tag{60}$$

The root-mean-square value of the electric field, averaged with respect to time and with respect to polar coordinate over a sphere of that radius r which goes with the dimensionless coordinate x is,

$$E_{\text{rms}}=\{-8\pi[T\varphi]\}^{\frac{1}{2}}=\{c^4f^2e^{-\nu}/Gr^2\}^{\frac{1}{2}}$$
$$=c^3\Omega G^{-\frac{1}{2}}b^{-\frac{1}{2}}l^{*-\frac{1}{2}}\varphi(x)$$
$$=(c^7/\mathcal{J}G^3)^{\frac{1}{2}}l^{*-1/6}[(1-b^2)^{\frac{1}{2}}/2b]\varphi(x)$$

with a peak value

$$E_{\text{peak rms}}=0.46l^{*\frac{1}{2}}c^4/G^{\frac{1}{2}}M.$$

4. TRANSFORMATIONS AND INTERACTIONS OF PURE ELECTROMAGNETIC GEONS

Energy leaks out of a simple idealized spherical geon at a rate easily estimated by the methods of the theory of alpha decay. Were the refractive index barrier removed, the energy would disappear from its present region of concentration in a time of the order of one vibration period. The attrition, α, in the equation of definition

$$dM/M=-\alpha\omega dt, \tag{61}$$

would have a value of the order of unity. Owing to the presence of the barrier, the attrition is cut down to a value of the order

$$\alpha \sim \exp(-2P), \tag{62}$$

where P is the barrier penetration integral determined by Eq. (38):

$$P = \int [(d^2 f/d^* \rho^2)/f]^{\frac{1}{2}} d\rho^*$$

$$= \int [-1 + (l^* Q/\rho)^2 (1 - 2L/\rho)]^{\frac{1}{2}} \times Q^{-1}(1-2L/\rho)^{-1} d\rho. \tag{63}$$

The integral extends from the outer edge of the active zone to the point of reemergence from the barrier. The relevant range in ρ is of the order of $l^* = [l(l+1)]^{\frac{1}{2}}$, whereas the thickness of the active zone is only of the order $\Delta\rho \sim l^{*\frac{1}{2}}$. For this reason it is legitimate to use for Q and L in the evaluation of the integral (63) the constant values that apply outside the active zone. The error made in the penetration exponent by this approximation will be only of the order of $l^{*\frac{1}{2}}$, whereas the exponent itself will be of the order l^*. Thus in accordance with the results at the end of the last section we write

$$Q = 1,$$
$$\Omega r_{\text{inner}} = \rho_{\text{lower limit}} = bl^* \doteq 0.33 l^*,$$
$$L = (b/2)(1-b^2) l^* \doteq 0.15 l^*, \tag{64}$$

and find by algebraic examination of the roots of the bracket in (63) that the outer turning point is

$$\Omega r_{\text{outer}} = \rho_{\text{upper limit}} = l^* \{ -(b/2) + [1 - (3b^2/4)]^{\frac{1}{2}} \} \equiv Bl^* \doteq 0.79 l^*. \tag{65}$$

The penetration integral has the value

$$P = l^* \int_b^B \{x^{-2}[1-(1-b^2)(b/x)] - 1\}^{\frac{1}{2}} \times [1-(1-b^2)(b/x)]^{-1} dx$$
$$= 0.76 l^*, \tag{66}$$

where the variable of integration is $x = \rho/l^* = \Omega r/l^*$; and the attrition has the rough value

$$\alpha \sim \exp(-1.52 l^*). \tag{67}$$

As examples, consider two simple idealized spherical geons not very far—on a logarithmic scale—above the limit where quantum effects come in by way of pair production. Let both have masses of 10^{42} gram and only one-half wave of disturbance in the radial direction in the active zone, but let the wavelength, and therefore the thickness of this zone, have quite different values in the two cases, as indicated in Table I. It is evident from the numbers given there that geons, even systems of the same mass and radius, can have fantastically different rates of radiation leakage.

To switch to quantum language, we can speak of the leakage process as single photon emission. In addition to such processes there will occur what can temporarily describe as double photon processes (Fig. 5): two quanta moving tangenially collide and go off in two new directions, not far from parallel or antiparallel to the radius vector, and thus escape from the system simultaneously. To follow the terminology of light rays a little further, we can speak as is well known, of a critical angle required for escape. For rays whose angle of inclination, θ, to the radius vector is greater than this critical amount, there exists a maximum distance to which the ray can go before it falls back on the geon. This maximum distance is obtained by calculating the appropriate root of the equation

$$\frac{l^{*2} \sin^2\theta}{\Omega^2 r^2}\left[1 - \frac{l^* b(1-b^2)}{\Omega r}\right] - 1 = 0. \tag{68}$$

The quantity on the left-hand side of this equation evaluated for the case $\sin\theta = 1$, came into the argument of the barrier penetration integral, another illustration of the close connection between the Hamilton-Jacobi analysis of rays and the differential equation for waves. It provides a simple interpretation of expression (68) to think of $l^* \sin\theta$ as a measure of the angular momentum of the ray: either a ray on the inside trying to get out; or a ray coming from infinity and trying to get in. The turning points for these two cases undergo merger when we have zero not only for the left-hand side of (68), but also for its derivative with respect to r. This happens when

$$\sin\theta = (3^{\frac{1}{2}}/2) b(1-b^2) \doteq 0.77 = \sin 50.4°, \tag{69}$$

and the double root then lies at

$$r = (3l^*/2\Omega) b(1-b^2) \doteq 0.44 l^*/\Omega, \tag{70}$$

TABLE I. Leakage of radiation from simple idealized spherical geons: illustrative examples. The last column gives the factor of change of scale for a similarity transformation which leaves the geon a classical object.

	Geon I	Geon II	Scale factor
Mass	10^{42} g	10^{42} g	$\times n$
Radial coordinate of active zone (58)	1.67×10^{14} cm	1.67×10^{14} cm	$\times n$
Spherical harmonic index $l^* = [l(l+1)]^{\frac{1}{2}}$	10	8.43×10^{9}	$\times 1$
Circular frequency of emergent radiation (57)	6.00×10^{-4} rad/sec	5.06×10^{1} rad/sec	$\times (1/n)$
Wavelength outside	3.14×10^{14} cm	3.72×10^{5} cm	$\times n$
Approximate attrition (67)	2.5×10^{-7}	$10^{-8570000000}$	$\times 1$
Time to collapse assuming leakage only and classical behavior throughout	212 years	$\sim \infty$	$\times (1/n)$
Rms electric field in most active region [Eq. (60) ff]	4.66×10^{19} esu/cm	4.41×10^{13} esu/cm	$\times (1/n)$
Critical field for pair production	4.41×10^{13} esu/cm	4.41×10^{13} esu/cm	a constant

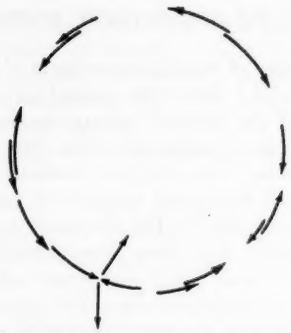


FIG. 5. Photon-photon collisions provide one of the mechanisms for escape of energy from the active region of a geon. The same mechanism is describable from the wave point of view in terms of a coupling between otherwise independent modes of oscillation of the electromagnetic field. The two sources of the nonlinear coupling are virtual pair phenomena and the nonstatic parts of the gravitational field.

that is, at the peak of the barrier in Fig. 2. This distance should be compared with impact parameter, or distance of closest approach in the absence of gravitational forces, of a photon with the same energy and angular momentum:

$$\text{impact parameter} = l^* \sin\theta/\Omega = 0.77 l^*/\Omega. \quad (71)$$

In contrast, the rays that hold the geon together have in this sense an impact parameter equal to $1.00 l^*/\Omega$, and have the inner and outer turning points,

$$r \doteqdot 0.33 l^*/\Omega \quad \text{and} \quad r \doteqdot 0.79 l^*/\Omega. \quad (72)$$

Rays with an impact parameter less than $0.77\ l^*/\Omega$ have a component of motion in the radial direction sufficiently great to pass freely from the inner region to the outer region.

Everything just stated in ray language can be rephrased more appropriately in terms of waves. In the quasi-static gravitational field of force of the geon we have already found a solution of the wave equation for the electromagnetic potential which has only one maximum in the active region, varies with angle as a spherical harmonic of order $l \doteqdot l^* - \frac{1}{2}$, has a proper circular frequency $c\Omega$. In addition to waves of such a kind excited to a strength sufficient to hold the geon together, we can have proper solutions of weak amplitude with a great variety of values for the number, s, of radial maxima and minima in the active region and of values of the indices l and m of the spherical harmonic. A typical disturbance of this type will be characterized by a proper value, $\Omega_{s,l}$, and an amplitude factor $A_{s,l,m}$. Let the wavelength of this mode be small compared to the dimensions of the geon. Then it is appropriate to characterize this mode, too, by an equivalent ray-optical impact parameter, $P \equiv \{(l/\Omega_{s,l})$ times a quantity having the significance of $\sin\theta$ and to be calculated via appropriate analysis from the relative values of s and $l\}$. The motion of the trapped ray will have an outer bound equal to the next to the largest solution of the equation

$$(P^2/r^2)[1-(l^*b/\Omega r)(1-b^2)]-1=0. \quad (73)$$

Outside of this distance the strength of the mode in question falls off exponentially through a refractive index barrier. A standing wave of this kind is the proper transcription of the idea of photon with a radial component of motion too small to allow escape. Evidently the prohibition of escape is not absolute. There will be leakage through the barrier to an extent the greater, the closer P falls to the critical impact parameter for escape, $P=0.77 l^*/\Omega$. For smaller values of the impact parameter we have no proper solutions with an exponential region of fall off. Instead, we have solutions capable of transporting energy freely from the inside of the geon to outer space, and endowed with a continuum of frequency values. These waves correspond to the notion of photons able to escape from the system.

Instead of speaking of photon-photon collisions, we can talk of excitation of secondary waves by the waves that carry the primary energy supply of the geon. Such excitation is not envisaged in a linear wave equation, derived from a Lagrangian function that contains the fields to the second order. The differentiation of a quadratic Lagrangian gives a time rate of change of the field strength proportional only to the first power of field quantities. However, the Lagrangian actually contains correction terms of the fourth order and higher. They give in the expression for the time rate of change of a field quantity supplementary terms containing the product of three or more field quantities. These terms constitute a nonlinear coupling that takes energy slowly out of the primary modes and redistributes it over secondary radiations. Some of this energy is picked up in characteristic modes. In any such mode the energy is reflected back and forth in a standing wave limited by two values of r. Other parts of the energy go into unbounded waves and are lost to infinity. The existence of the barrier leakage phenomenon means of course that the distinction between the stationary modes and the unbounded waves is only approximate. The stationary modes will themselves leak some of the energy that they get. Moreover these vibrations can transfer energy via nonlinear couplings to still other modes, and also to still other waves that escape.

The relative importance of simple leakage, and of transfer of energy to free-running waves, depends upon the relative strength of excitation of the various modes of the geon. The detailed specification of the state of a general geon is so complicated that there exists the greatest variety of objects, showing amongst them the greatest extremes of behavior.

In a geon where nonlinear terms make the more important contribution to the energy dissipation, we evidently face a problem so complicated that statistical arguments are needed to make any headway. We have to take into account the totality of the stationary modes of the geon in an approximately static gravita-

tional field, and the distribution of energy among these modes. But their number is infinite, as in Rayleigh's paradox of blackbody radiation. No proper account of the distribution can be given without taking into account the quantum of action. We are therefore invited to assign to each mode of characteristic frequency ω an energy $\hbar\omega[\exp(\hbar\omega/\Theta-1]$, where Θ, expressed in energy units, has the significance of a temperature. We can also idealize the energy of each unbounded wave as zero in a first approximation, these waves being most easily distinguished by the criterion (73). The very number of the modes of vibration simplifies the problem. To find in each mode the radial distribution of field strength, and stress and energy, it is no longer necessary to integrate the wave equation numerically. The JWKB method gives the answer in terms of purely algebraic operations. Thus the system of three equations of the theory of spherical geons reduces to two differential equations of the first order for the metric quantities $\lambda(r)$ and $\nu(r)$, or $L(r)$ and $Q(r)$, with the temperature Θ as parameter of integration. No attempt is made here to solve this pair of equations for the "thermal geon."

The thermal geon is obviously an idealization that suggests itself for following out the consequences of the nonlinear interactions between different modes of vibration. In the next approximation one has to make further allowance for the effects of the intermodal couplings, by way of radiation losses through photon-photon collisions. Here the picture will be very different according as the mean free path of the most relevant photons is large or small in comparison with the dimensions of the system. In the second case it becomes necessary to analyze the effective opacity of the system, and to take account of the variation of temperature with radial coordinate in a fashion familiar from stellar theory—a problem again not investigated here.

The nonlinear couplings arise from two sources: the electron pair field, and the gravitational field. The second is classical. The source of the gravitation field—the stress energy tensor—has a reasonably smoothly varying average value, but on top of this are superposed fluctuations due to the fact that the electromagnetic field does after all vary in space and in time. A typical constituent of the fluctuation will have a character qualitatively of the form $A_1 f_1(x,y,z)\sin(\omega_1 t+g_1)\ A_2 f_2(x,y,z)\times\sin(\omega_2 t+g_2)$. This term will give rise to a fluctuating component of the gravitational field. Consequently in the differential equation for the electromagnetic field the coefficients of metric origin will not be exactly static functions of position alone, but will have small alternating terms with circular frequencies, $(\omega_1+\omega_2)$ and $(\omega_1-\omega_2)$, and with amplitudes proportional to A_1A_2. Let the uncorrected amplitude of the mode under consideration be $A_3 f_3(x,y,z)\sin(\omega_3 t+g_3)$. Then the wave equation for this oscillation has to be visualized as containing four supplementary source terms, with frequencies $\omega_1\pm\omega_2\pm\omega_3$, and with amplitudes proportional to the product $A_1A_2A_3$. Thus the primary disturbances in the geon inevitably generate secondary waves.

For an order of magnitude estimate of the strength of the effect, let λ denote the general scale of the space variations in the relevant primary modes, and let $\mathcal{E}$ denote the order of magnitude of the associated electromagnetic fields. Then a typical inhomogeneity in the distribution of stress and energy will possess a mass of the order of $\mathcal{E}^2\lambda^3/c^2$. The fluctuations in the gravitational metric will have the general magnitude $\Delta g\sim G\Delta m/c^2\lambda\sim(G/c^4)\lambda^2\mathcal{E}^2$. The square of this dimensionless factor, multiplied by a dimensionless function of the geometry of the geon—depending upon ratio of typical wavelengths to the size of the system, and upon other details—will determine the fraction of the energy of the system lost per cycle via nonlinear coupling processes of gravitational origin.

The correction effects due to virtual production of electron pairs by slowly varying but intense fields bring in a coupling which has the same qualitative consequences as that due to gravitational fields, except that the governing dimensionless factor, Δg, of the previous paragraph is to be replaced by $\sim(e^2/\hbar c)[\mathcal{E}e(\hbar/mc)/(mc^2)]^2$. Thus the pair effects depend upon the electric field itself, whereas the gravitational effects depend upon the electric potential, $A\sim\lambda\mathcal{E}$. The two effects become of the same order of magnitude for wavelengths of the order

$$\lambda\sim(e^2/Gm^2)(G\hbar/c^3)^{\frac{1}{2}}$$
$$=(4.16\times10^{42})(1.60\times10^{-33}\text{ cm})=0.67\times10^{10}\text{ cm}. \quad (74)$$

To justify the classical analysis of geon structure, we already know that we have to deal with *radii* of the order of the limit (74) or greater. On the wavelengths we have not previously had any limit, except inferiority to the radius. However, in dealing with geons in which many modes are excited, or especially with a thermal geon, it will ordinarily be reasonable to consider wavelengths quite small compared to the limit (74). Then for most purposes gravitational transfer of energy between modes can be neglected in comparison with the coupling due to the charges and currents of virtual pairs.

In analyzing the behavior of a geon classical in the sense of being large compared to the limit (74), we encounter both a large scale static gravitational field and small scale rapid variations in this average. There will exist in addition long scale periodic or secular changes in the configuration of the system, some of which will in certain cases describe significant mechanisms for the disintegration of a geon. It is easy to visualize a slow vibration in the case of a simple toroidal geon. The shape of the torus changes from circular to elliptical first in one sense and then in the other, with the elementary electromagnetic disturbances, or ray tracks if one will, adjusting themselves adiabatically to these slow readjustments in the distribution of refractive index. The gravitational interaction between the various

elementary electromagnetic oscillations imposes upon their readjustments of shape a collective character. In these collective motions the coupling between the electromagnetic oscillators by way of gravitational interactions will exceed the coupling via virtual pair effects, so long as we are in the classical domain of geon sizes.

The same mode of vibration of a simple toroidal system, endowed with enough energy, will lead to deformations not far in form from a figure eight, and with still further excitation of this mode scission will occur. Whether the two separate rings then completely break their association depends upon the magnitude of the kinetic energy of recoil from the act of scission. When this energy exceeds the gravitation potential energy of attraction between the two objects, they fly apart to make two distinct systems, and we have a complete act of fission. When the recoil energy is less than the binding, the two rings separate to a limiting distance, reverse their motion, collide, pass through each other, again separate, and so on. At each act of collision some of the energy of relative motion will be redistributed. Some will go into excitation of the electromagnetic modes of the one ring, some into modes of the other ring, some into modes that owe their existence to the combined gravitational field of the pair of rings, some into collective vibrations and rotations of the individual rings, and some will escape into space as free radiation. Ultimately the relative motion of the two tori will be damped down and they will come together to form a single geon with a smaller mass, and a much more complicated distribution of energy among modes, than characterized the original system.

Fission can in principle take place spontaneously in any classical geon. The principal requirement is adequate time for the exchange of energy between the various modes of the system, so that ultimately, by a statistical fluctuation in the distribution of energy, enough becomes concentrated upon a collective mode of distortion to lead to a critical deformation followed by fission. As between the mechanisms of energy dissipation, and compared to radiation leakage and nonlinear coupling to free running waves, the probability of fission will vary in a complicated way depending upon whether the distribution of energy among modes is finegrained or coursegrained, whether the gravitational field is symmetric or has large scale irregularities, and whether the total angular momentum is small or large. These circumstances will also affect the distribution in size of the fragments from fission in those classes of cases where this process occurs with appreciable probability.

A geon is characterized not only by its internal structure and by the genetics of its radioactive decay processes, but also by its interaction with other geons. When the two systems pass by each other at a separation large compared with the size of either of them, they will act on each other like the bodies of classical physics. The interaction is the relativistic generalization of the simple Newtonian law of force, $F=-GM_1M_2/r_{12}$.[2] The electromagnetic interaction between the two systerms will give rise to an additional force which is enormously weaker than the gravitational force. Thus, the flux of outgoing leakage radiation from one of the systems undergoes scattering by the other geon and gives rise to a radiation force of the order

$$F\sim(-c^2dM_1/cdt)[\pi R_2^2/4\pi r_{12}^2]$$
$$\sim(cM_1\alpha_1\omega_1)[R_2(GM_2/c^2)/r_{12}^2]$$
$$\sim(GM_1M_2/r_{12}^2)(\alpha_1 l_1 R_2/R_1). \quad (75)$$

An additional force will arise from the pressure of the radiation emitted by the other body. In expression (75) the attrition, α_1, is an exponentially small function of the order, l_1; thus the dimensionless product $\alpha_1 l_1$ will be extremely small for all stable geons. Of course, if the two geons are exceedingly unlike in size, or one of them is decaying so fast as to constitute an explosion, then the radiation force between the objects can become comparable to the gravitational force.

When two geons have large intrinsic angular momenta, the forces depend upon orientation as well as upon distance. Let two simple toroidal geons be oriented with their principal planes perpendicular to the line that connects their centers. When the rays of light in one geon all go around one way, and those in the second go around the other way, then the attraction is stronger than it is in the case of parallel orientation of the two angular momentum vectors. This circumstance is a reminder that a gravitational field is described by quantities more complicated than a static potential. The supplementary orientation-dependent forces constitute in this case a fraction $\sim R_1R_2/r_{12}^2$ of the total gravitational force.

As one geon flies by another at the minimum separation r_{12} and with a relative velocity v perpendicular to this line, it not only sets up in the other geon a bulk gravitational field of the order G_1M_1/r_{12}^2, but also creates tide producing forces of the order $GM_1R_2^2/r_{12}^4$. Both forces have frequency components ranging from $\omega=0$ to $\omega\sim v/r_{12}$. The first acts on the center of mass, a degree of freedom with zero natural frequency, and produces a bulk motion in accord with the classical concept of "body." The second acts on a mode of deformation with a natural frequency of the order c/R_2, very much higher than that of the driving force. Consequently the response will be adiabatic and a negligible part of the energy of motion will be left in internal degrees of freedom after the two geons have gone far apart.

When two geons pass by each other at a distance which brings the outer boundaries of their active regions to a separation, ΔR, small in comparison with geon radii, then the interaction between the two objects becomes more complicated. First, on account of the tidal deformations the gravitational forces will experience fractional increases of the order R_2^2/r_{12}^2 and R_1^2/r_{12}^2.

Second, there will come into being fluctuating forces of electromagnetic origin, of the order of magnitude of the product of typical field strengths, $\mathcal{E}_1$ and $\mathcal{E}_2$, in the active regions of the two geons, multiplied by an exponential decrement factor of the character $\exp(-\Delta R/\lambda\!\!\!^{-})$, multiplied by an area of contact factor of the order $R\lambda\!\!\!^{-}$. Here we have assumed that the important wavelengths in the two geons are of the same general magnitude. Also we have used R to represent the radius of the smaller of the two geons; or better, R can be identified with the reduced radius, $R_1R_2/(R_1+R_2)$. Let f_1 denote the fraction of the volume of the first geon occupied by its active region, so that

$$\mathcal{E}_1^2 f_1 R_1^3/c^2 \sim M_1 \sim c^2 R_1/G. \tag{76}$$

Then the electric field in the active region of this geon will be of the order

$$\mathcal{E}_1 \sim f_1^{-\frac{1}{2}} G^{\frac{1}{2}} M_1/R_1^2, \tag{77}$$

and the expression for the electromagnetic force between the two systems will have the character

$$F \sim (f_1^{-\frac{1}{2}} f_2^{-\frac{1}{2}} \lambda\!\!\!^{-}/R)(GM_1M_2/r_{12}^2) \exp(-\Delta R/\lambda\!\!\!^{-}). \tag{78}$$

As the first dimensionless combination of factors will be of the order of unity for many classes of geons, it follows that the interaction (78) is of the order of the gravitational force, multiplied by a characteristic exponential function of distance.

Two geons which almost touch as they pass will display not only an anomalous interaction but also energy exchanges and energy losses. Where the refractive index barrier between the two systems is thinnest, radiation will leak across from one body to the other. Some modes previously excited weakly or not at all will gain energy, and other important modes will lose energy. Thus the two bodies will part company with a relation of masses different from what they had when they met. Their total mass will be diminished at the same time, both via momentary stimulation of the leakage of radiation to the outside, and by way of local perturbations of the gravitational field that deflect some light rays to the outside from orbits that previously were relatively stable.

Let two geons collide still more directly, so that they pass through each other. It will be rare for the two systems to emerge from the encounter without much loss of mass or redistribution of mass. The frequency spectrum of the perturbations experienced by one of the geons during the encounter will range from tidal forces of frequency $\sim c/R$ at one end of the spectrum, through strong components of frequency $\sim c/$(thickness of active region), to fluctuations with the scale of wavelengths and frequencies associated with the individual electromagnetic modes themselves. As a result almost everything that can happen ordinarily will happen. There will be a substantial loss of energy in the form of free radiation. Collective modes of motion will be excited. Individual electromagnetic modes will have quite new amplitudes. Moreover, forces will be at work to induce fission of the original geons into smaller fragments. In view of all these circumstances, it is not appropriate to describe the close interaction of two classical geons in terms of an effective law of force, or potential curve. Instead, one has to use the much more complicated language of transmutation physics, and to try to describe the distribution in mass of the products of a geon reaction. Evidently the geon decay processes and interaction mechanisms form an extensive subject for a more detailed investigation, not undertaken here.

5. INFLUENCE OF VIRTUAL PAIRS ON GEON STRUCTURE

As we go down in the mass and size spectrum of geons, we eventually pass out of the purely classical realm (Region I) into conditions (Region II) where the field strengths are no longer negligible compared to the critical field, $\mathcal{E}_{\text{crit}}=m^2c^3/\hbar e$, of the theory of electron pairs; and from here it is but a short step to Region III, where the fields are strong enough to turn a problem of virtual pairs into one where real pairs are present in large numbers. About III we say nothing; and about II our considerations are very primitive. The induced charges and currents of the virtual pairs alter the electromagnetic properties of the medium, so that it acquires a refractive index, n. It will be reasonable to consider all the electromagnetic modes of the geons under consideration to have wavelengths very long in comparison with the Compton length, $\hbar/mc$. Under these conditions the refractive index will be independent of frequency but will be dependent upon field strength, $\mathcal{E}$:

$$n-1 \sim (e^2/\hbar c)(\mathcal{E}/\mathcal{E}_{\text{crit}})^2. \tag{79}$$

To be more specific, let the geon be a simple toroid of azimuthal index number a. Then the relevant field strength in (79) is some appropriate average over the active region. The speed of light around the torus is reduced to c/n. The connection between radius of the geon and frequency of its leakage radiation will be altered to the form $\omega \sim nac/R$, where the integer, a, represents as before the azimuthal index number. The general relation between changes in energy and changes in reduced action, $\mathcal{J}=\text{action}/2\pi$, becomes

$$dM=c^{-2}\omega d\mathcal{J} \sim nad\mathcal{J}/cR=[1+a(\mathcal{J}_{\text{crit}}/\mathcal{J})]ad\mathcal{J}/cR. \tag{80}$$

In the last line for convenience we have translated expression (79) for the refractive index correction from a field strength dependence to an action dependence, according to relation (5) between the two variables, the uncorrected formula (5) being legitimate in the evaluation of a first order correction. The deflection of the radiation in the gravitational field is more difficult to discuss. In lieu of an accurate analysis, for which only the bare formalistic bones are presented below, the following order of magnitude analysis is presented, which may well be in error through oversight of some significant factor. A photon which outside has a fre-

quency ω has inside a mass of the order $\hbar\omega/c^2$, and experiences in the gravitational field of the geon a force of attraction of the order $(GM/R^2)(\hbar\omega/c^2)$. This expression should equal the time rate of change of the momentum of the photon: the product of its momentum, $n\hbar\omega/c$, by the angular rate of revolution in its orbit, c/nR; that is, $\hbar\omega/R$. Equating the gravitational and dynamic terms gives the same relation that we had in the absence of a refractive index correction:

$$M\sim(c^2R/G)(?).$$

Multiplying this formula by (80), we have

$$MdM\sim(acd\mathcal{J}/G)[1+a(\mathcal{J}_{\rm crit}/\mathcal{J})](?). \qquad (81)$$

Here the question mark indicates that the relation in question is subject to possible correction. If this relation is correct, it states that the square of the geon mass, though nearly a linear function of action for values of the action large in comparison with a certain characteristic value, nevertheless for lower values of the action curves towards a steeper dependence on $\mathcal{J}$. The region where (81) should apply is too small on the scale of Fig. 6 to allow showing this curvature there.

The dependence of effective refractive index increment upon the square of the field strength lets itself be seen in two different ways.[12] From the photon point of view the torus is a channel filled by two streams of quanta moving in opposite directions. The space density of photons going one way is of the order

$$n_1\sim\mathcal{E}^2/\hbar\omega_1. \qquad (82)$$

The photons travelling the other way are not able to make pairs by the process[13] $\hbar\omega_2+\hbar\omega_1\rightarrow e^++e^-$ through want of energy: $\hbar\omega_1=\hbar\omega_2\ll mc^2$; but could do so were their quantum energies greater than $\hbar\omega^*=(mc^2)^2/\hbar\omega_1$. For this process the absorption cross section starts at zero at the threshold $\hbar\omega_2=\hbar\omega^*$, rises to a peak value $\sigma\sim(e^2/mc^2)^2$ at a small multiple of this frequency, and then falls off. The absorption presented by the medium at high frequencies implies a contribution to the refractive index, n, at low frequencies, according to the formula

$$\begin{aligned} n(\omega_2)-1&=(c/\pi)\int_{\omega^*}^{\infty}n_1\sigma(\omega_2')d\omega_2'/(\omega_2'^2-\omega_2^2)\\ &\sim(c\mathcal{E}^2/\hbar\omega_1)(e^2/mc^2)^2/\omega^*\\ &\sim(e^2/\hbar c)[\mathcal{E}e(\hbar/mc)/(mc^2)]^2, \end{aligned} \qquad (83)$$

in agreement with (79). Alternatively stated, there exists a process of scattering of light by light,[6] for which the differential cross section at low frequencies in the forward direction,

$$(d\sigma/d\Omega)_{\rm forward}\sim(e^2/\hbar c)^2(e^2/mc^2)^2(\hbar\omega_2/mc^2)^4(\hbar\omega_1/mc^2)^2, \qquad (84)$$

makes a contribution to the refractive index given by Rayleigh's formula,

$$n-1=n_1 2\pi(c/\omega_2)^2(d\sigma/d\Omega)^{\frac{1}{2}}_{\rm forward}; \qquad (85)$$

this contribution is in agreement with (79).

This analysis of the nonlinear behavior of a strong electromagnetic field makes it clear that the corrections to be applied in Region II will be very different for a toroidal geon according as it has an angular momentum of zero, so that half the photons go each way, or has the maximum possible angular momentum, so that the photons meet no counterstream. In the second case the onset of strong pair corrections will evidently be postponed to higher field strengths and smaller bodies.

The picture of photon-photon collisions and a refractive index is not suitable for a precise evaluation of the

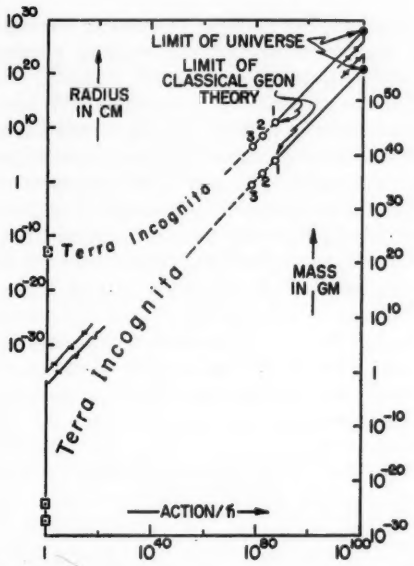


FIG. 6. Mass and radius as functions of action for classical geons (slanting full lines in upper right hand portion of diagram). With diminishing mass and radius, the electric fields active in an electromagnetic geon increase to a point, 1, where electron pairs are produced, below which point a classical analysis of geons completely fails (dot-dash line in diagram). For pure neutrino geons it may be possible to carry a simple analysis down as far as point 2 in the diagram before pair effects appear (Sec. 6 of text). Point 3 indicates where densities of the nuclear order of magnitude would be attained in either electromagnetic or neutrino geons were extrapolation of simple geon theory justified, which it certainly is not. Neither is any further extrapolation of the curves for mass and radius at all allowed (*Terra Incognita*). Masses and radii of electron and proton are indicated on the diagram for orientation as to magnitudes. The representation of geon mass or radius by a single line, or the action variables of the system by a single number, gives a rather oversimplified picture, as is clear from Eqs. (5) for even the simplest variety of toroidal geons. The numbers in the present diagram may be thought of as applying to simple toroidal geons with azimuthal index number, a, equal to 10 or less.

[12] John S. Toll, Ph.D. thesis, Princeton University, 1952 (unpublished); J. S. Toll (to be published); J. S. Toll and J. A. Wheeler, Phys. Rev. **81**, 654 (1951) and more detailed account, to be published.

[13] G. Breit and J. A. Wheeler, Phys. Rev. **46**, 1087 (1934).

pair corrections. For example, in a simple toroidal geon of zero angular momentum the standing electromagnetic wave with which we deal will, if strong, produce changes in the properties of the medium with twice the periodicity of the wave itself. This phase relation means that no simple averaging of the properties of the medium is appropriate. For a detailed treatment of the corrections it is therefore appropriate to go back to first principles, as typified by the variational principle (4) of classical relativity theory. Heisenberg and Euler[14] have analyzed the case of slowly varying but strong fields and have shown that the charges and currents of the virtual pairs have the same effect as if there were no pairs, but instead the Lagrangian of the variation principle of the electrodynamics of special relativity theory were corrected from $\mathfrak{L}$, where

$$-\mathfrak{L}=(\mathbf{E}^2-\mathbf{B}^2)/8\pi=(1/8\pi)\mathcal{E}_{\text{crit}}{}^2S, \tag{86}$$

to $\mathfrak{L}^*$, where

$$(-\mathfrak{L}^*)-(-\mathfrak{L})=(1/8\pi^2)(e^2/\hbar c)\mathcal{E}_{\text{crit}}{}^2\int_0^\infty dx e^{-x}x^{-3}$$

$$\times\left\{ix^2P\frac{\cos[x(S+2iP)^{\frac{1}{2}}]+\text{conj.}}{\cos[x(S-2iP)^{\frac{1}{2}}]-\text{conj.}}-(x^2S/3)+1\right\}. \tag{87}$$

Here S represents the scalar, $(\mathbf{E}^2-\mathbf{B}^2)/\mathcal{E}_{\text{crit}}{}^2$ and P represents the pseudoscalar, $(\mathbf{E}\cdot\mathbf{B})/\mathcal{E}_{\text{crit}}{}^2$. For small field strengths,

$$-8\pi\mathfrak{L}^*/\mathcal{E}_{\text{crit}}{}^2=S+\pi^{-1}(e^2/\hbar c)$$
$$\times\{(S^2+7P^2)/45+(26P^2S+4S^3)/315+\cdots\}. \tag{88}$$

In general relativity the appropriate scalar and pseudoscalar quantities are

$$\mathcal{E}_{\text{crit}}{}^2S=-\tfrac{1}{2}F_{\alpha\beta}g^{\alpha\gamma}g^{\beta\delta}F_{\gamma\delta},$$
$$\mathcal{E}_{\text{crit}}{}^2P=\tfrac{1}{8}(-g)^{\frac{1}{2}}\epsilon^{\alpha\beta\gamma\delta}F_{\alpha\beta}F_{\gamma\delta}. \tag{89}$$

Here g represents the determinant of the g_{ik}. The quantities $\epsilon^{\alpha\beta\gamma\delta}$ do not form a tensor. They are defined instead by the statements that $\epsilon^{1234}=1$ and that ϵ changes sign on the interchange of any two indices. To correct for the effect of virtual pairs in geon theory we have only to replace $\mathfrak{L}/c=(1/16\pi c)\ F_{\alpha\beta}F^{\alpha\beta}$ in the variational principle (4) of general relativity by $\mathfrak{L}^*/c$, calculated as just defined. In taking the variation of the integral, which includes the volume factor $(-g)^{\frac{1}{2}}dx^1dx^2dx^3dx^4$, we regard as the quantities to be varied the 16 gravitational potentials g^{ik} and the four electromagnetic potentials A_i in $F_{ik}=(\partial A_k/\partial x^i)-(\partial A_i/\partial x^k)$. The coefficient of δg^{ik} contains two parts, of which one is the Einstein analog, $G_{ik}\equiv R_{ik}-\frac{1}{2}g_{ik}R$, of the d'Alembertian of the gravitational potential, and the other is the stress-energy tensor, T_{ik}:

$$T_{ik}=2(-g)^{-\frac{1}{2}}\delta[(-g)^{\frac{1}{2}}\mathfrak{L}^*]/\delta g^{ik}$$
$$=2(\partial\mathfrak{L}^*/\partial S)(\delta S/\delta g^{ik})$$
$$+2(\partial\mathfrak{L}^*/\partial P)(\delta P/\delta g^{ik})-g_{ik}\mathfrak{L}^*$$
$$=-2F_{i\beta}g^{\beta\delta}F_{k\delta}\mathcal{E}_{\text{crit}}{}^{-2}(\partial\mathfrak{L}^*/\partial S)$$
$$-g_{ik}(P\partial\mathfrak{L}^*/\partial P+\mathfrak{L}^*). \tag{90}$$

Thus the gravitational field equations take their usual form, $G_{ik}=(8\pi G/c^4)T_{ik}$, with only the change (90) in the form of the stress energy tensor. Similarly the coefficient of the variation, δA_i, of a typical electromagnetic potential, equated to zero, gives the ith electromagnetic field equation,

$$0=4\pi(-g)^{-\frac{1}{2}}(\partial/\partial x^\alpha)[(-g)^{\frac{1}{2}}\partial\mathfrak{L}^*/\partial(\partial A_i/\partial x^\alpha)]$$
$$=4\pi(-g)^{-\frac{1}{2}}(\partial/\partial x^\alpha)\{(-g)^{\frac{1}{2}}(\partial\mathfrak{L}^*/\partial S)$$
$$\times(2F^{i\alpha}/\mathcal{E}_{\text{crit}}{}^2)+(-g/2)(\mathcal{E}_{\text{crit}}{}^{-2}\partial\mathfrak{L}^*/\partial P)\epsilon^{\alpha i\gamma\delta}F_{\gamma\delta}\}, \tag{91}$$

where the F's are considered to be expressed in terms of the A's. Equations (90) and (91) define the theory of geons in Region II.

NEUTRINO-CONTAINING GEONS

There is little difference between the theory of a geon built out of neutrinos, and one built out of electromagnetic fields, apart from the fact that each neutrino state will accommodate only one quantum of energy. The general relativity version of Maxwell's equations is replaced by the general relativity form of Dirac's equation[15] for an entity of zero rest mass. The stress-energy tensor for Dirac particles of zero rest mass forms the source term in the gravitational field equations. The equation $GM/c^2R\sim 1$ still connects mass and radius of the object. The modes of dissipation of energy are still leakage through the refractive index barrier, coupling through nonlinear effects to waves that run freely to infinity, and various forms of fission. Consequently there is little point to reformulating in terms of neutrinos our discussion of the properties of electromagnetic geons.

One difference appears in the process of interpenetration of two geons. When both are made of neutrinos, the overlap of the active regions of the two objects will force a promotion of their constituent neutrinos to states of higher momentum and energy, in accordance with the Pauli principle. The uptake of energy implies a strong effective repulsion compared with the forces that would otherwise have been at work, as for example in the collision of an electromagnetic geon and a neutrino geon, or the impact of a neutrino geon on an anti-

[14] W. Heisenberg and H. Euler, Z. Physik **98**, 714 (1936); see also V. Weisskopf, Kgl. Danske Videnskab. Selskab., Mat.-fys. Medd. **14**, No. 6 (1936–7).

[15] See for example Marcel Riesz, "L'équation de Dirac en relativité générale" (to be published). I am indebted to Professor Riesz for the opportunity to see this paper in advance of publication.

neutrino geon. Thus there is a certain interesting specificity about the interaction of geons.

A second and fundamental difference shows between pure neutrino geons and pure electromagnetic geons when one asks how small the object can be before quantum effects come in. In the neutrino case quantum effects are of course in principle present right from the start, in the action of the Pauli exclusion principle. But no other quantum effects are evident, and one does not at first sight see any reason why one should not be able to follow the properties of neutrino geons down to very small sizes, enormously less than the limit $\sim 10^{10}$ cm placed on classical electromagnetic geons by the onset of virtual pair phenomena. However, one at least knows where the limit is in the electromagnetic case, while in the neutrino case the critical limit could be anywhere, as indicated by the following analysis. Just as the nonlinear phenomena of electromagnetic geons commence when photon-photon collisions become important, so neutrino-neutrino collisions represent a potential critical mechanism to set a lower bound on the size, and an upper limit on the density, of those pure neutrino geons that are susceptible to simple analysis. The penetrating power of a neutrino through ordinary matter is so enormous[16] that it seems at first sight ridiculous to consider the collision of neutrino with neutrinos. However, if the continuous spectrum of electrons from μ-meson decay is correctly interpreted in terms of the simultaneous emission of two neutrinos (or a neutrino and an antineutrino) then it necessarily follows that two neutrinos of the appropriate character running towards each other with sufficiently high energy, $(\hbar\omega_1)\cdot(\hbar\omega_2) \gtrsim (207mc^2/2)^2$, will necessarily be capable of provoking the reaction $\nu+\nu \rightarrow \mu^-+e^+$; and from the Fermi type of beta-decay theory it follows that the reaction cross section will vary at sufficiently high energies as

$$\sigma \sim (g^2/\hbar^2c^4)\omega_1\omega_2, \tag{92}$$

where the coupling constant is of the order

$$g \sim 10^{-49} \text{ erg cm}^3. \tag{93}$$

The existence of this absorption mechanism for high energies implies that even low-energy neutrinos can produce virtual μ,e pairs, which reannihilate and send the neutrinos off into new directions. The cross section for this neutrino-neutrino scattering process, evaluated in the forward direction by use of the dispersion relation[12]

$$(d\sigma/d\Omega_2)_{\text{forward}} = \left| (\omega_2^2/2\pi^2c) \int_0^\infty \sigma_{\text{absn}}(\omega_2')(\omega_2'^2-\omega_2^2)^{-1} d\omega_2' \right|^2, \tag{94}$$

is easily seen to diverge logarithmically,

$$(d\sigma/d\Omega_2)_{\text{forward}} \sim (\omega_1\omega_2^2 g^2/\hbar^2c^5)^2 \ln^2(\hbar\omega_{\text{upper}}/100\, mc^2), \tag{95}$$

unless one inserts for the indicated upper limit of infinity a finite upper limit, $\hbar\omega_{\text{upper}}$, for which one so far has no evidence. Similarly, the refractive index presented by a medium containing per unit volume n_1 neutrinos of circular frequency ω_1 differs from unity, according to (83), by the amount

$$n(\omega_1)-1 \sim (g^2/\hbar^3c^3)(n_1\hbar\omega_1) \ln(\hbar\omega_{\text{upper}}/100\, mc^2), \tag{96}$$

again a divergent expression.

In view of the divergence of the neutrino-neutrino interaction as evaluated from Fermi beta theory and the dispersion relation, we really have no basis to discuss neutrino geons at all, much less to set a lower limit of masses at which the theory takes a new form, as in the case of electromagnetic geons. Nevertheless, it is still of interest to see what we can say if we assume that the logarithm, instead of diverging, has a value of the order of 10^2 or less. In this case we can employ (96) to draw two conclusions. First, it will be beyond the scope of simple neutrino theory to analyze geons in which the refractive index correction is of the order of unity, or the energy density exceeds

$$\begin{aligned} n_1\hbar\omega_1 &\sim (\hbar^3c^3/g^2)[1/\ln(\hbar\omega_{\text{upper}}/100\, mc^2)] \\ &\sim 10^{47} \text{ erg/cm}^3 \\ &\sim (10^{26} \text{ g/cm}^3)c^2. \end{aligned} \tag{97}$$

To yield such a density, neutrino states must be occupied up to momenta of the order p, or energies of the order cp, where

$$10^{47} \text{ erg/cm}^3 \sim \int_0^p (cp)2(4\pi p^2 dp/h^3), \tag{98}$$

or

$$\begin{aligned} cp &\sim (\hbar^3c^3 10^{47} \text{ erg/cm}^3)^{\frac{1}{4}} \\ &\sim 1 \text{ erg} \sim 10^6\, mc^2. \end{aligned} \tag{99}$$

This energy is so great that it carries us beyond the domain of virtual μ,e processes to real production of pairs of this kind. This circumstance leads to the second conclusion, that the energy density of any geon which can be analyzed in terms of simple neutrino theory can at most be equal to the right-hand side of (98), evaluated for a maximum neutrino energy, cp, of the threshold value, $100\, mc^2$:

$$\begin{aligned} \text{energy density} &\sim 2\pi(cp)^4/\hbar^3c^3 \\ &= 3.62\times 10^{31} \text{ erg/cm}^3 \\ &= (4.02\times 10^{10} \text{ g/cm}^3)c^2, \end{aligned} \tag{100}$$

corresponding according to (96) to a refractive index correction of only one part in 10^{15}. This estimate implies a mass density considerably less than a typical estimate for nuclear matter, $\sim 1.4\times 10^{14}$ g/cm^3. It is conceivable that there is some limitation of which physics is not

[16] See, for example, F. Reines and C. Cowan, Phys. Rev. **92**, 830 (1953).

aware that puts a bound to the density of pure neutrino geons much smaller than (100). However, if real production of μ,e pairs sets the only limit, then it is possible in principle to treat in terms of existing theory objects built up solely out of neutrinos and gravitational forces down to a radius

$$R=[(M/R)(3/4\pi)(4\pi R^3/3M)]^{\frac{1}{2}}$$
$$\sim[(c^2/G)1.69\times10^{11}\ \text{g/cm}^2)]^{\frac{1}{2}}$$
$$=2.83\times10^{8}\ \text{cm}$$

and mass

$$M\sim(c^2/G)R=3.82\times10^{36}\ \text{g}. \qquad (101)$$

This lower limit for pure neutrino geons is less than the corresponding limit (10) for pure electromagnetic geons (10^{11} cm, 10^{39} g) by not more than a few powers of ten. In one case the bound comes from the limiting electromagnetic field strength, or field energy density, at which pairs of positive and negative electrons begin to appear; in the other case, electron-mu-meson pairs.

Mixed geons, energized by a combination of neutrinos and electromagnetic fields, will have properties similar to either kind of pure system, and hardly need separate consideration here. However, the presence of neutrinos and photons on an equal footing does raise again the question of the neutrino theory of light, developed by Jordan, Kronig, and others, and brought to a halt by the discovery of Pryce that the theory in its then existing form did not possess proper relativistic invariance.[17] The theory assumes that there exists between neutrinos a physical interaction, never introduced explicitly and never discussed, such that a photon is described by a pair consisting of one neutrino and one antineutrino; or rather, by a quantum-mechanical linear superposition of very many such pairs of states. The description implies an unavoidable complementarity, such that a statement of the occupation numbers of the photon states, and a prescription of the occupation numbers of the neutrino states, stand to each other in a mutually exclusive relation. When only neutrino states are occupied, the theory speaks of a pure neutrino field. When in addition a few antineutrino states are filled, a mixture of light and neutrinos is said to be present. With equal numbers of neutrinos and antineutrinos a pure photon field is considered to exist; and so on, up to the case of a pure antineutrino field. Such a description would evidently subsume geons of purely electromagnetic character, of purely neutrino type, and of mixed constitution, all into a unified class of systems.

Regarding the status of the neutrino theory of light, I am kindly informed by Professor Eugene Wigner that the conceivable mechanisms for the combination of the spinors and momentum vectors of the neutrino and antineutrino states to form in an invariant way the vector magnitudes of the photon states are far from having been explored in a comprehensive way in the literature, so that is not necessarily clear that the objections of Pryce will forever retain their force. Second, recent studies of the decay of the μ meson[18] show that the lifetime and the form of the electron spectrum[19] are together consistent with a universal Fermi interaction of the kind met in beta decay only if the two neutrinos are of opposite character. This result suggests, though it does not prove, that μ decay produces one neutrino and one antineutrino. The consequences are the existence of the $\nu+\nu^*\rightarrow\mu+e$ reaction at high energies, and ν,ν^* scattering at low energies. In other words, a physical background for the interaction of the neutrino theory of light does not have to be postulated; it exists. Third, no such argument exists for an interaction between two neutrinos of the same character. Consequently it is conceivable that the theory of pure neutrino geons, or pure antineutrino geons, can be carried without meeting new physical effects to energy densities larger, and sizes smaller, than the limits of (100) and (101).

In summary, if we assume the existence of a cutoff in the Fermi interaction at high energies in Eqs. (95) and (96), then the door is open to analyzing the properties of neutrino geons over an enormous range of sizes without going outside the scope of existing theory; but below an uncertain critical limit of sizes interesting and fundamental physical questions raise themselves.

7. ELECTRICITY, GAUSS' THEOREM, AND GRAVITATIONAL FIELD FLUCTUATIONS

What of free electricity? Consistently to complete the scheme of classical physics we find no alternative but to regard the geon as exemplar of the concept of body; but neither in this object nor in the divergence free theory behind it is there any place for charge. All lines of force continue without end. Let a sphere be drawn around the immediate neighborhood of a point. Then as many lines of force go in as out.

In applying the theorem of Gauss we have tacitly assumed space to be simply connected. One knows, however, that the notion of Riemannian manifold by itself places no such requirement upon the space-time continuum. One can consider a metric which on the whole is nearly flat except in two widely separated regions, where a double-connectedness comes into evidence as symbolized in Fig. 7. The general divergence-free electromagnetic disturbance holding sway in the space around one of these "tunnel mouths" will send forth lines of force into the surrounding space, and appear to have a charge. However, an equal number of lines of force must enter the region of disturbance from

[17] P. Jordan, Ergeb. exakt. Naturwiss. **7**, 158 (1928); Z. Physik, numerous papers ending with **105**, 229 (1937); R. de L. Kronig, several papers ending with Z. Physik **100**, 569 (1936) and Physica **3**, 1120 (1936); M. H. L. Pryce, Nature **141**, 976 (1938). See also de Broglie, Heisenberg, and Kramers in *L. de Broglie, Physicien et Penseur* (Editions Albin Michel, Paris, 1953). I am indebted to Professor Arthur Wightman for several discussions of the present status of Jordan's idea.

[18] L. Michel and A. S. Wightman, Phys. Rev. **93**, 354 (1954).

[19] J. Vilain and R. W. Williams, Phys. Rev. **92**, 1586 (1953).

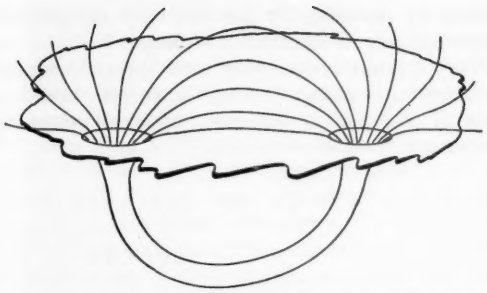


FIG. 7. Schematic representation of lines of force in a doubly-connected space. In the upper continuum the lines of force behave much as if the tunnel mouths were the seats of equal and opposite charges.

the tunnel. Consequently the other mouth of the tunnel must manifest an equal and opposite charge. In such a doubly-connected space it is evidently a matter of definition whether we say that divergence-free field equations do or do not permit the existence of electric charge. It will be convenient to say yes if the width of the tunnel is small compared to the separation of its mouths. So far we have inquired only after the behavior of the electromagnetic field in a metric assumed to be pre-existing. However, in classical relativity theory the metric cannot be taken arbitrarily, but must be found by solution of the gravitational field equations. No investigation is known to have been made of the possibility of a self-consistent solution that is double-connected. Yet one would not be surprised to find that no reasonable choice of boundary conditions would permit such a classical solution.

Let one pass from the classical theory defined by the action principle (4) to the corresponding quantum theory, either by the prescription of Feynman,[20]

$$\psi\sim \sum_{\substack{\text{all conceiv-}\\ \text{able relevant}\\ \text{field histories}}} \exp[(i/\hbar)\ (\text{classical action for each field history})], \qquad (102)$$

or by any other standard method. Then one is forced to recognize the existence of fluctuations in all fields. Their magnitudes depend upon the size, L, of the space-cotime regions under consideration, and are given under suitable conditions by formulas of the form $\Delta F\sim(\hbar c)^{\frac{1}{2}}/L^2$; $\Delta A\sim(\hbar c)^{\frac{1}{2}}/L$; and $\Delta g\sim(\hbar G/c^3)^{\frac{1}{2}}/L$. So long as one deals with distances large in comparison with $(\hbar G/c^3)^{\frac{1}{2}}\sim10^{-33}$ cm, one can disregard for most purposes the fluctuations in the metric, and consider space to be simply connected. But in deriving Gauss' theorem one is driven to consider an integral over the whole of the region in question, including regions of the very smallest spatial extension. Here the inevitable fluctuations force on space time a most complicated structure. Because it is the essence of quantum mechanics that *all* field histories contribute to the probability amplitude, the sum (102) not only may contain doubly and multiply connected metrics; it must do so. General relativity, quantized, leaves no escape from topological complexities of which Fig. 7 is only an oversimplified symbol. In this sense the door is open for the existence of charges in the quantum version of a theory that contains no charges.

Little progress has so far been made in studying the quantization of general relativity.[21] (1) It is not yet certain whether the method of summing over configurations gives in the case of nonlinear theories results that are identical to those derived from other methods of quantization. (2) It is not certain that the action function of general relativity can be given a well-defined meaning for those field configurations, classically unrealizable, that make the factor $(-g)^{\frac{1}{2}}$ in the action function a pure imaginary. This square root recalls the similar factor in relativistic electron theory, $(-dx_\mu dx^\mu)^{\frac{1}{2}}$, where likewise similar difficulties of interpretation arise for nonclassical paths, and where the root has been replaced by Dirac's matrix expression, $\Gamma_\mu dx^\mu$, associated with half-integral spins and Fermi-Dirac statistics. Thus, as of today we cannot exclude that charges will show themselves naturally and inevitably in the spinor quantization of the only comprehensive and divergence-free classical theory of fields that we possess.

8. CONCLUSION

Taking seriously and following out the consequences of the forty year old theory of general relativity, we have been led to recognize the relative stability of certain types of electromagnetic field disturbances held together by gravitational forces. These geons furnish for the first time a completely classical, divergence-free, self-consistent picture of the Newtonian concept of body over the range of masses from $\sim10^{39}$ g to $\sim10^{57}$ g. Two such geons interacting at a distance large compared to their characteristic dimensions behave as elementary objects. However, when one geon is followed for a long time or is allowed to interact closely with another it undergoes interesting and characteristic transformation processes.

Classical geons are not objects for study in the laboratory, nor is there any evident reason to believe that geons of the classical range of sizes now exist in nature, or ever did exist. Even were such large geons once present, a sufficient lapse of time would guarantee the decay of all but extraordinarily stable ones to a mass below the limit where these systems are capable of classical analysis.

On entry into the quantum domain of sizes, electromagnetic geons build up field strengths strong enough to produce pairs of electrons. At not very different sizes

[20] R. P. Feynman, Phys. Rev. **76**, 769 (1949).

[21] I am indebted to Professor James Anderson and Professor Peter Bergmann for instructive discussions of the literature of this problem.

neutrino geons commence to give birth to μ meson-electron pairs. Consequently the projection of geon theory to objects of dimensions less than $\sim 10^{10}$ cm to $\sim 10^{8}$ cm requires the analysis of new phenomena.

In conclusion, the geon makes only this visible contribution to science: it completes the scheme of classical physics by providing for the first time an acceptable classical theory of the concept of body.

One's interest in following geon theory down into the quantum domain will depend upon one's considered view of the relation between very small geons and elementary particles.

PHYSICAL REVIEW VOLUME 97, NUMBER 2 JANUARY 15, 1955

Angular Momentum of a Real Field

R. H. DICKE
Palmer Physical Laboratory, Princeton University, Princeton, New Jersey
(Received May 11, 1954; revised manuscript received October 7, 1954)

It is paradoxical that a field oscillating in a mode which on grounds of symmetry contains no angular momentum should carry angular momentum when quantized. This angular momentum is shown to result from zero-point oscillations in other modes. The operator for the square of the angular momentum of the field is discussed.

THE angular momentum of a second quantized real field has been discussed a number of times in the past in connection with liquid drop oscillation[1] and multipole expansions of the electromagnetic field.[2] These articles discuss the z component of the angular momentum in a satisfactory fashion but the discussion of the square of the angular momentum seems to have been first carried out in a proper way only recently.[3] It would hardly seem necessary to raise again the question of the angular momentum of a second quantized field, but there seems to be considerable confusion in the literature as regards to the square of the angular momentum. Also in at least three of these references[2] there are remarks which are not quite correct. The fact that there is a good deal of evidence that this old but important question is still not properly understood has convinced the author that someone should write still another note on this subject even if it does not contain very much that is new.

In order to illustrate the type of question that we wish to consider, we discuss first a liquid drop oscillating in a surface mode given by the spherical harmonic $l=2$, $m=0$, namely an oscillation back and forth between prolate and oblate spheroidal shapes. It is obvious from symmetry arguments that such a mode of oscillation should carry no angular momentum. How then in second quantization for a one-quantum state of this mode, does one expect to get a squared angular momentum of $6\hbar^2$ rather than zero? The usual answer which one encounters is that this result of $l=2$ for the quantum number of squared angular momentum of the excited mode follows in an elementary way from the fact that the excited mode is described by a spherical wave with that index. This, however, is not in accord with the facts. In classical field theory the square of the angular momentum is given by the square of the m index of the spherical mode which is excited. It might be expected that after the inclusion of quantum fluctuation effects the angular momentum of the field for a one particle state would be $m(m+1)\hbar^2$. However, in second quantization for a one-particle state it becomes a function of the l index rather than the m index for reasons which are quite subtle. The easiest way to get the correct result is to remember that an alternative description of the quantized field is a first quantized many-particle system where for one particle the square of the angular momentum is, of course, given by the l index of the ψ field. However, in field theory the problem is much less transparent.

To come back again to the example given above, the symmetry arguments (reflections in x, y, and z planes) which show that classically there should be zero angular momentum associated with this mode, somehow become invalid when the system is described in quantum-mechanical language. Since each of the normal modes of oscillation constitutes a separate dynamical system quite independent of the others, it is clear that their symmetry properties exist independent of their level of oscillation or of possible zero-point fluctuations in other modes, and that the angular momentum of this mode should, in fact, be zero. If this seems parodoxical, it is, in part, because of a confusion in terminology. For a classical field it is possible to speak of the angular momentum of a normal mode of oscillation in an

[1] M. Fierz, Helv. Phys. Acta **16**, 365 (1943).
[2] W. Heitler, Proc. Cambridge Phil. Soc. **32**, 112 (1936); W. Heitler, *Quantum Theory of Radiation* (University Press, Oxford, 1949); J. B. Blatt and V. Weisskopf, *Theoretical Nuclear Physics* (John Wiley and Sons, Inc., New York, 1952); B. Stech, Z. Naturforsch. **7a**, 401 (1952); R. G. Sachs and J. G. Brennan, Phys. Rev. **88**, 825 (1952).
[3] C. M. DeWitt and J. H. D. Jensen, Naturforsch. **8a**, 267 (1953). It is proposed to consider here questions not covered in this reference.

unambiguous way. However, in a quantum-mechanical description, because of the zero-point fluctuations in other modes, this becomes unambiguous only if the angular momentum is a sum of the angular momentum of the separate modes of oscillation of the field. Because the z component of the angular momentum is always[1,2] shown to be the sum of the angular momenta of the separate modes of the field expansion, whereas little is said about the other components, one could easily incorrectly assume that this is also true of the other components. It is easily shown, however, and was shown in the first paper on the electromagnetic problem[2] that the other two components contain cross terms between modes of the same l index and frequency. Because of this, the angular momentum of the field is not the sum of the angular momenta of the separate modes used to describe the field. In other words it is quite meaningless, in the case of a quantized field, to speak of the angular momentum of a mode of the field. One must always speak of the angular momentum of the total field which depends in a complex way on the oscillations of all the modes of the field.

The argument based on symmetry was improper when applied to the quantum field for a simple reason. Although each of the normal modes of the field may have a certain symmetry property, the total field oscillation generally does not possess this property when more than one mode is excited.

A simple example may serve to illustrate this point. A three-dimensional oscillator, in the form of a particle oscillating about the origin, has zero angular momentum about the origin when the particle oscillates along either the x or y axis. Furthermore, this zero angular momentum results from the symmetry of these normal modes of oscillation. However, with both modes present simultaneously and a 90° phase displacement between them, the oscillator has a nonzero z component of angular momentum.

There are still a number of interesting questions which cannot be answered by such simple qualitative arguments as those given above. For example, granted that zero-point oscillations contribute to the angular momentum, why does the zero particle state of the field have no angular momentum? What can be said about the squared angular momentum with more than one particle present? Is it well defined for states of definite occupation number? What are the simultaneous eigenstates of energy and angular momentum? Instead of giving a partial answer to these questions for the total field, the question will be simplified somewhat, then answered in a fairly complete way. Inasmuch as the angular momentum cross terms occur between modes of the same frequency and l index, the interesting physical questions are seen even if the field excitation is limited to such a set of modes. This considerably simplifies the problem as there are no zero-point contributions from other modes and they can be ignored.

Although the following remarks are limited to a scalar field φ, a real transverse vector field can be generated from φ as $iL^{-1}\mathbf{L}\varphi$ or $ic\int_{-\infty}^{t}\mathbf{curl}L^{-1}\mathbf{L}\varphi dt'$.[4] Also a longitudinal vector field can be generated as $c\int_{-\infty}^{t}\mathbf{grad}\varphi dt'$. Here $\mathbf{L}$ is the usual angular momentum operator, L is the positive square root of $(\mathbf{L})^2$ and c is the wave velocity. It can be shown that the usual expressions for the angular momentum of the resulting vector fields, when integrated over a spherical enclosure as described below, is identical with that of the scalar field used to generate them. Also the energy of the scalar field, excluding its $l=0$ part, is equivalent to the energy of the vector field. Consequently, the following arguments will apply also to electromagnetic fields (where the field variable is a vector potential) and to sound fields. As an example of the scalar field itself, $\partial\varphi/\partial t$ can be taken proportional to the pressure change of a sound wave in a gas.

It should be emphasized that although the ratio of the field angular momentum to energy for the vector field is identical with that of the generator scalar field, excluding the $l=0$ part, this equivalence holds only for the integrated expressions and not for the densities.

A real scalar φ satisfying the wave equation

$$\Delta\varphi-\frac{1}{c^2}\frac{\partial^2\varphi}{\partial t^2}=0 \tag{1}$$

is assumed to have an energy density

$$U=\tfrac{1}{2}\left(\frac{\partial\varphi}{\partial t}\right)^2+\tfrac{1}{2}c^2(\mathbf{grad}\varphi)^2, \tag{2}$$

a momentum density

$$\mathbf{p}=-\frac{\partial\varphi}{\partial t}\mathbf{grad}\varphi, \tag{3}$$

equal to the energy flux density divided by the wave velocity squared, and an angular momentum density

$$\mathbf{r}\times\mathbf{p}=-\frac{i}{\hbar}\frac{\partial\varphi}{\partial t}\mathbf{L}\varphi. \tag{4}$$

The field is assumed confined to a sphere, appropriate boundary conditions being satisfied on its surface. The energy, momentum, and angular momentum of the field are integrals of Eqs. (2), (3), and (4) over the interior of the sphere.

If φ is expanded in any orthonormal set of functions, the angular momentum of the field when expressed as a function of the expansion coefficients must contain cross products between the various modes. This follows

[4] It has been shown that $\mathbf{L}\varphi$ represents a transverse vector field [see W. Franz, Z. Physik **127**, 363 (1950)]. For the present purposes, it is necessary to include the operator L^{-1} to properly normalize individual terms in a spherical wave expansion and, consequently, to make the following result valid for a general field distribution.

from the noncommutation of the components of the operator $\mathbf{L}$. If the expansion is in the usual spherical harmonics, the z component of the field angular momentum contains no cross terms but the x and y components consist wholly of products of expansion coefficients of modes having the same frequency and l index but differing in m index by ± 1 unit.

For an expansion in real harmonics, all three components contain only cross terms. As stated previously, the cross terms are between modes of the same frequency and l index.

Since in quantizing the field the operators for the energy, momentum, and angular momentum of the total field are integrals over the interior of the sphere of the corresponding densities given by Eqs. (2), (3), and (4), the angular momentum operator is

$$[\mathbf{L}]=-\frac{i}{\hbar}\int\frac{\partial\varphi}{\partial t}\mathbf{L}\varphi d\tau, \tag{5}$$

where the bracket is used to signify an operator used in second quantization of the field. Without the bracket the symbol signifies the usual single particle operator from first quantization. It might be assumed that the operator for the squared angular momentum is also of this form, namely,

$$\text{“}[L^2]\text{”}=-\frac{i}{\hbar}\int\frac{\partial\varphi}{\partial t}L^2\varphi d\tau. \tag{6}$$

This, however, is incorrect. The resulting operator is devoid of physical interest. It is not Hermitian and has a Hermitian part which is ambiguous.

Of much greater interest is another closely related operator. Let us express φ in terms of a complex field as

$$\varphi=\psi+\psi^*. \tag{7}$$

Here ψ is assumed to be the negative-frequency part of φ. Expressed in terms of these field quantities, the operators for energy, momentum, and angular momentum of the total field can all be written as the following integral over the interior of the enclosure:

$$[Q]=-\frac{2i}{\hbar}\int\frac{\partial\psi^*}{\partial t}Q\psi d\tau, \tag{8}$$

where Q is the usual one-particle operator from first quantization, namely,

$$Q=-\frac{\hbar}{i}\frac{\partial}{\partial t},\quad \frac{\hbar}{i}\mathbf{grad},\quad \text{and}\quad \frac{\hbar}{i}\mathbf{r}\times\mathbf{grad}, \tag{9}$$

respectively. Note that Eq. (8) is closely related to the expectation value of a first quantized single-particle theory. One might attempt to construct an operator for the square of the angular momentum by setting $Q=L^2$. The resulting operator is Hermitian, and does commute with both the Hamiltonian of the field and the angular momentum operators. It also has correct eigenfunctions and eigenvalues as long as there is no more than one quantum in the field. For more than one quantum, its eigenvalues and eigenfunctions are almost all incorrect. The incorrectness of this operator follows from the fact that in field theory a squared angular momentum density is meaningless and there is no expression for the square of the angular momentum in the form of an integrated density.

The correct operator for the square of the angular momentum of the field is, of course

$$[L^2]=[\mathbf{L}]^2=[L_x]^2+[L_y]^2+[L_z]^2. \tag{10}$$

In order to avoid a great deal of complication, which does not contribute to an understanding of the problem, these operators will only be expressed explicitly for a development of the field in $l=1$ real spherical harmonics. Since the angular momentum operator, Eq. (5), contains cross terms only between modes of the same frequency and l index, most of the important physical results can be seen most easily by ignoring all of the modes of the field except those of some particular frequency and l index, for which the simplest case is $l=1$.

By expanding in real harmonics, all three modes of the field are characterized by a symmetry for which their angular momentum is classically zero. Thus, all three components of the angular momentum of the total field contain only cross terms between these modes. This expansion does not give the z axis a special significance and all three components are treated in a symmetrical manner. This leads to certain simplification over an expansion in the usual spherical harmonics based on a rotation axis.

Expand φ as

$$\varphi=\sum_j q_j Y_j(\theta,\varphi)g(r). \tag{11}$$

Here j takes on the three values x, y, and z. Y_x is the $l=1$, $m=0$ spherical harmonic with x as its polar axis. q_x is assumed real. The energy of the field may be expressed in terms of the expansion coefficients and conjugate momenta $p_j=\dot{q}_j$ as

$$[H]=\tfrac{1}{2}\sum_j(p_j^2+\omega^2q_j^2)=\sum_j a_j^*a_j+\tfrac{3}{2}\hbar\omega$$
$$=\sum_j(n_j+\tfrac{1}{2})\hbar\omega, \tag{12}$$

where $a_j=(p_j-i\omega q_j)/\sqrt{2}$ is an annihilation operator and n_j is the occupation number of the j mode. The angular momentum of the field has a z component [from Eq. (5)]:

$$[L_z]=-\frac{i}{\hbar}\int\frac{\partial\varphi}{\partial t}L_z\varphi d\tau=q_xp_y-q_yp_x$$
$$=\frac{1}{i\omega}(a_x^*a_y-a_z^*a_z), \tag{13}$$

with similar relations for the x and y components.

The formal similarity of this operator with that for the angular momentum of a particle results from the transformation properties of these spherical harmonics under coordinate system rotations.

It should be noted that, because of the presence of the annihilation operators, $[L_z]$ has the eigenvalue zero for any state of excitation of the z oscillator, the x and y oscillators being unexcited. Note, also, that if either the x or y oscillator is excited, but not both, the state is not an eigenstate of $[L_z]$ but does have an expectation value of zero consistent with the classical result of zero expected from symmetry arguments. Also, as is expected, all three components of $[\mathbf{L}]$ have an eigenvalue zero for the zero-particle state of the field.[3]

In terms of a development of the field in these normal modes, the expression for the square of the angular momentum is

$$[L^2]=2\hbar^2[(n_x+\tfrac{1}{2})(n_y+\tfrac{1}{2}) + (n_y+\tfrac{1}{2})(n_z+\tfrac{1}{2})+(n_z+\tfrac{1}{2})(n_x+\tfrac{1}{2})-\tfrac{3}{4}] - (1/\omega^2)[a^{*2}a^2-\textstyle\sum_j a_j^{*2}a_j^2]. \quad (14)$$

Here the first term in brackets on the right-hand side is a function only of the occupation number operators. In the second term in brackets,

$$a^2=\mathbf{a}^2=a_x^2+a_y^2+a_z^2. \quad (15)$$

Note the effect of the zero-point oscillations in causing the half-integers to appear in the first bracket. The second bracket has an eigenvalue of zero for any state of the field for which occupation numbers are well defined but no occupation number exceeds unity, and consequently such states are eigenstates of $[L^2]$. For such states, the first bracket of Eq. (14) has the eigenvalue $n(n+1)\hbar^2$, where $n=n_x+n_y+n_z$. There are eight such states. For all other states of definite occupation number L^2 is not well defined. Its expectation value for such states is, however, given by the first bracket of Eq. (14). The eigenvalue for the zero- and one-particle states are in agreement with DeWitt and Jensen.[3] However, the remaining four states of definite occupation number appear only in this formalism and have their origin in the use of real spherical harmonics.

To find all the simultaneous eigenstates of $[L_z]$, $[L^2]$ and the energy, it is only necessary to note that from Eqs. (12) and (13), the Hamiltonian and angular momentum operators are formally identical with those of a three-dimensional oscillator, and known results can be used in the description of the field. In particular, for an n-particle state of the field, possible total angular momentum quantum numbers are $l=n, n-2, n-4, \cdots$.

All the simultaneous (unnormalized) eigenfunctions of $[H]$, $[L_z]$, and $[L^2]$ can be generated as

$$\psi_{nlm}=([L_z]-i[L_y])^{l-m}(a_x^*+ia_y^*)^l(a^{*2})^{(n-l)/2}\psi_{000}.$$

This follows from the commutation relations of $\mathbf{a}^*$ with $[\mathbf{L}]$.[5]

Since the only states of definite occupation number for which L_z is well defined are those for which only the z mode is excited, and in view of the fact that L^2 is well defined for only eight states of definite occupation number, there are only two of the above states [Eq. (14)] which are simultaneously states of definite occupation number, namely

$$\psi_{000} \quad \text{and} \quad \psi_{110}.$$

It is interesting to note that the expectation value of L^2 for large occupation number of one normal mode, the others being unexcited, is zero in agreement with the classical result. Here the classical limit is to be taken by letting the occupation number go to infinity, keeping its product with $\hbar$ finite.

It is a pleasure to acknowledge helpful discussion with many of my Princeton colleagues and with Professor V. Weisskopf of the Massachusetts Institute of Technology.

[5] E. U. Condon and G. H. Shortley, *The Theory of Atomic Spectra* (Cambridge University Press, London, 1935), pp. 56–63. $\mathbf{a}^*$ falls in the class of vectors designated as $\mathbf{T}$. The first operator $(a^{*2})^{(n-l)/2}$ creates $n-l$ particles without giving the field angular momentum. The next operator creates l particles and l units of angular momentum. The last operator, as usual, reduces L_z to $m\hbar$. That this set of states is complete can be seen by considering the resulting degeneracies.

Goldstein's Eigenvalue Problem

H. S. GREEN
University of Adelaide, Adelaide, South Australia
(Received August 26, 1954)

A mathematical error is found in Goldstein's solution of the Salpeter-Bethe equation with zero total energy and a cut-off factor. When this error is corrected, no solution remains of the eigenvalue problem.

I. INTRODUCTION

ALL solutions of the Salpeter-Bethe equation so far obtained for the two-nucleon system have been based on rather drastic approximations whose quantitative effect is not easy to see. Klein[1] has examined rather exhaustively the limitations of the "ladder approximation"; but he allowed two other common approximations, the neglect of the time of propagation of the virtual mesons, and the omission of pair contributions, to go unchecked. Arnowitt and Gasiorowicz[2] have recently discussed pair effects, but the effect of the "static potential" approximation is still largely unknown, and most likely could only appear from a completely covariant solution of the S-B equation.

For this reason, the covariant treatment of the problem by Goldstein[3] was of great interest, even though it was limited to the ladder approximation and a highly unrealistic value (zero) of the total energy. Goldstein showed that the S-B equation as it stands, at least in the special case he considered, has a solution for any value of the coupling constant. He suggested that the only way to obtain an eigenvalue of the coupling constant, consistent with the occurrence of a bound state of the two nucleons, was to introduce a cut-off factor in the kernel of the S-B equation. Unfortunately, as will appear below, his solution of the resulting eigenvalue problem has a mathematical flaw, the elimination of which invalidates his conclusion.

Solutions of the S-B equation exist of a more general type than, but for the same total energy as, that which Goldstein considered; but all solutions investigated disappear when a cutoff is introduced. The reason is probably related to the extreme value of the binding energy assumed. So Goldstein's essential idea—that the bound-state problem should be solved by a limiting process involving the use of a convergence factor—may very well be right in general. For independent reasons stated elsewhere,[4] the author does in fact agree with Goldstein on the matter. It is hoped that approximate covariant solutions of the S-B equations for general values of the binding energy will soon be available to afford a more conclusive test of the idea.

[1] A. Klein, Phys. Rev. **90**, 1101 (1953); **91**, 740 (1953); **92**, 1017 (1953); **94**, 1052 (1954).
[2] R. Arnowitt and S. Gasiorowicz, Phys. Rev. **94**, 1057 (1954).
[3] J. S. Goldstein, Phys. Rev. **91**, 1516 (1953).
[4] I. E. McCarthy and H. S. Green, Proc. Phys. Soc. (London) **A67**, 719 (1954).

II. SOLUTION OF THE S-B EQUATION

The integral equation whose solution is required is

$$(\not{p}-m)\varphi(p)(\not{p}+m)=\frac{\lambda}{i\pi^2}\int\frac{c(k^2,\Omega)\gamma_5\varphi(k)\gamma_5 d^4k}{(p-k)^2}, \quad (1)$$

where $c(k^2,\Omega)$ is the cut-off factor which has a value near unity for $k^2\ll\Omega$ and near zero for $k^2\gg\Omega$. Goldstein proposed to solve it by reducing it to a differential equation; this can be achieved, without making any special assumptions about the form of $\varphi(p)$, by applying the operator $\Box=\partial^2/\partial p^\mu\partial p_\mu$ to both sides of (1). Since

$$\Box\{(p-k)^2\}^{-1}=-(2\pi)^2i\delta(p-k), \quad (2)$$

the result is

$$\Box\psi(p)+4\lambda c(p^2,\Omega)\gamma_5\varphi(p)\gamma_5,$$
$$\psi(p)=(\not{p}-m)\varphi(p)(\not{p}+m). \quad (3)$$

This equation has to be supplemented by boundary conditions, which can be obtained by inspection from (1); these are

$$\psi(p)=\text{finite constant}, \quad p=0; \quad (4)$$

$$p^2\psi(p)=\text{finite constant}, \quad p^2\gg\Omega. \quad (5)$$

If now one assumes with Goldstein that $\varphi(p)$ has the form

$$\varphi(p)=\Phi(p^2/m^2), \quad (6)$$

Eq. (3) reduces to

$$X''(s)+\lambda C(s,\Omega)\Phi(s),$$
$$X(s)=s(s-1)\Phi(s), \quad (7)$$

with $s=p^2/m^2$. If one takes $C(s,\Omega)=1$ for $s<\Omega$ and $C(s,\Omega)=0$ for $s>\Omega$, the conditions (4) and (5) require that

$$\Phi(0)=\text{finite constant} \quad (8)$$

and

$$X'(\Omega)=0, \quad (9)$$

which are readily seen to be equivalent to Goldstein's conditions (16a) and (16b) (the latter with Ω substituted for ∞).

Goldstein gave

$$\Phi(s)=\text{const}\times F(1+\alpha, 2-\alpha; 2; s) \quad (10)$$

as that solution of (7) which satisfies (8), where the constant α is a root of the equation

$$\alpha(\alpha-1)+\lambda=0 \quad (11)$$

and is to be determined from the boundary condition (9). The "hypergeometric function" is, however, easily expressed in terms of Legendre functions, as one can see by writing

$$z=2s-1, \tag{12}$$

whereupon (7) becomes, for $z<\Omega$,

$$(z^2-1)X''(z)+\alpha(\alpha-1)X(z)=0. \tag{13}$$

Since

$$\frac{d}{dz}\left\{(z^2-1)\frac{dP_{\alpha-1}(z)}{dz}\right\}+\alpha(\alpha-1)P_{\alpha-1}(z)=0,$$

one sees that the general solution of (13) is

$$X(z)=\int_{-z}^{1}P_{\alpha-1}(z)dz+A\int_{1}^{z}P_{\alpha-1}(z)dz. \tag{14}$$

Here one must set $A=0$ to satisfy (8), and (9) reduces to

$$P_{\alpha-1}(1-2\Omega)=0. \tag{15}$$

For real positive values of λ, α is restricted to real values between 0 and 1, and complex values of the form $\frac{1}{2}(1+i\eta)$, where η is real. In the real domain of α, it is well known[5] that (15) can only be satisfied by values of Ω between 0 and 1. In the complex domain, (15) can only be satisfied by values of Ω less than 0. Hence (15) has no solution for values of Ω greater than unity.

To see how Goldstein obtained a different result, notice that the asymptotic form of (15) is

$$\frac{\Gamma(2\alpha-1)(-\Omega)^{-\alpha}}{\Gamma(\alpha-1)\Gamma(1-\alpha)}+\frac{\Gamma(1-2\alpha)(-\Omega)^{\alpha-1}}{\Gamma(\alpha)\Gamma(-\alpha)}=0, \tag{16}$$

or, since $\Gamma(z+1)=z\Gamma(z)$ and $\sin(\pi z)\Gamma(z)\Gamma(1-z)=\pi$,

$$\frac{\sin(\pi\alpha)}{\pi(2\alpha-1)}\{(\alpha-1)\Gamma(2\alpha)(-\Omega)^{-\alpha}+\alpha\Gamma(2-2\alpha)(-\Omega)^{\alpha-1}\}=0. \tag{17}$$

This agrees with the condition Goldstein at first obtains; but he proceeds to cancel a factor and write the condition in a form [Eq. (40)]:

$$\frac{(-1)^{2\alpha}\Gamma(2\alpha)}{\Gamma(2-2\alpha)}=\Omega^{2\alpha-1}\frac{\alpha}{\alpha-1},$$

which is satisfied by $\alpha=\frac{1}{2}$. The factor $(2\alpha-1)^{-1}$ in (17), however, may clearly *not* be removed for $\alpha=\frac{1}{2}$ without introducing a spurious solution; and the left-hand side of (17) in fact approaches a nonzero value as $\alpha\to\frac{1}{2}$.

III. OTHER TRIAL SOLUTIONS

One possible reason which suggests itself for the absence of a solution to Goldstein's eigenvalue problem, is that the bound-state solution might not be of the type assumed in (6). If one assumes instead that

$$\psi(p)=\Psi(p^2/m^2)\not{p}, \tag{18}$$

one obtains

$$(s-1)s\Psi''(s)+3(s-1)\Psi'(s)-\lambda\Psi(s). \tag{19}$$

The conditions (4) and (5) now require that $\Psi(0)$ should be finite and that the derivative of $s^2\Psi(s)$ should vanish for $s=\Omega$. [The latter is also the condition that $\Psi(s)$ and $\Psi'(s)$ should be continuous at $s=\Omega$.] The solution of (19) which is finite for $s=0$ is

$$\Psi(s)=F(a,b;3;s), \tag{20}$$

where

$$a=1+(1+\lambda)^{\frac{1}{2}},\quad b=1-(1+\lambda)^{\frac{1}{2}}, \tag{21}$$

and one has also

$$\frac{d}{ds}\{s^2\Psi(s)\}=2sF(a,b;2;s), \tag{22}$$

so that to satisfy the boundary condition at $s=\Omega$, $F(a,b;2;\Omega)$ must vanish. Now

$$F(a,b;2;\Omega)=\frac{\Gamma(a-b)(-\Omega)^{-a}}{\Gamma(b)\Gamma(a-2)}F(a,a-1;1+a-b;\Omega^{-1})+\frac{\Gamma(b-a)(-\Omega)^{-b}}{\Gamma(a)\Gamma(b-2)}F(b,b-1;1+b-a;\Omega^{-1}); \tag{23}$$

and since, for large values of Ω, Ω^a and Ω^b will be of different orders of magnitude, $F(a,b;2;\Omega)$ cannot vanish except possibly in the degenerate case when $a-b$ is an integer. Then $F(a,b;2,s)$ reduces to one of the Jacobi polynomials $F_n(2,2,s)$; but the zeros of these polynomials are known to lie in the interval $0<s<1$, so they cannot provide a solution of the problem either.

This investigation has shown that there are no covariant solutions of the integral equation (1) with real positive values of λ. There might possibly be noncovariant solutions, and since the derivation of Eq. (1) assumes a relativistic frame in which the mass-center of the two nucleons is at rest, it is fairly plausible that the bound-state solution, if any, should depend on $\not{p}_4$ and γ_4 as well as the covariant variables p^2 and $\not{p}$. But a full investigation shows that the independent solutions of (3), which are functions of $s=p^2$ and $z=\not{p}_4(p^2)^{-\frac{1}{2}}$, and are finite for $s=0$, are of the form $(s-1)^{-1}s^{\frac{1}{2}n}F(\frac{1}{2}n+\alpha,\frac{1}{2}n+1-\alpha;2+n;s)C_n^1(z)$, where α is the greater root of the equation

$$\alpha(1-\alpha)+\tfrac{1}{2}n(\tfrac{1}{2}n+1)=\lambda,$$

n is any integer, and $C_n^1(z)$ is the corresponding Gegenbauer polynomial. The boundary condition requires that the derivative $s^nF(\frac{1}{2}n+\alpha,\frac{1}{2}n+1-\alpha;n+1;s)$ of $s^{n+1}F(\frac{1}{2}n+\alpha,\frac{1}{2}n+1-\alpha;2+n;s)$ should vanish for $s=\Omega$,

[5] See W. Magnus and F. Oberhettinger, *Formulas and Theorems for the Special Functions of Mathematical Physics* (Chelsea Publishing Company, New York, 1949).

and this is impossible for any positive real value of λ. Adding a factor $\not{p}$ to the trial solution does not improve the situation.

Thus even the possibility of noncovariant solutions of the integral equation has to be excluded.

IV. DISCUSSION

The foregoing conclusions might be interpreted as casting doubt on the ability of the S-B equation to predict bound states. It might appear that the solutions of this equation obtained by various authors are merely a feature of the noncovariant approximations which they used. The present author, however, would prefer to take the view that the extreme value of the binding energy assumed by Goldstein is responsible for his failure to obtain a valid discrete value of the coupling constant. If one gave the binding energy of the ground state its maximum value (infinity) in the nonrelativistic approximation, one would not get a solution there either. One may thus maintain that to give a binding energy equal to the total rest-energy of the two nucleons, the coupling constant would have to be infinite, and that the possibility of a discrete finite value for any other binding energy is not excluded.

Goldstein stated that an expansion of the solution in powers of the total energy appeared to be singular, and if that is so it rather supports such a conclusion. But what is really needed is independent evidence of the nature of the solutions of the S-B equation for general values of the binding energy. Such evidence should not be founded on a noncovariant approximation, as covariance is clearly the crux of the matter. The author hopes to present a completely covariant treatment of the S-B equation in the near future.

PHYSICAL REVIEW VOLUME 97, NUMBER 2 JANUARY 15, 1955

Coulomb Functions for Large Charges and Small Velocities*

L. C. BIEDENHARN,† R. L. GLUCKSTERN, M. H. HULL, JR., AND G. BREIT
Yale University, New Haven, Connecticut

(Received July 30, 1954)

Expansions in powers of η^{-1}, where η is defined in the introduction below, for the Coulomb wave functions $F_L(\rho)$ and $G_L(\rho)$ and their derivatives are given for special values of $\rho=2\eta$ and $\rho=\rho_L=\eta+[\eta^2+L(L+1)]^{\frac{1}{2}}$, the classical turning points for $L=0$ and any L, respectively. Expansions applicable in the vicinity of the turning point are given as a series involving Bessel functions of order $\pm n/3$ with the expansion parameter ρ_L^{-1}. Approximations valid for large values of η are given and discussed.

I. INTRODUCTION

NUCLEAR reactions involving "heavy" charged particles[1,2] and the inelastic scattering of charged particles by nuclei[3,4] have recently been the object of several investigations, both theoretical and experimental. In both cases, the Coulomb interaction can be expected to play a dominant role, and the Coulomb wave functions are necessary for discussions of nuclear interactions of this type. It is evident, that for the parameter

$$\eta=ZZ'e^2/\hbar v,$$

which, together with $\rho=kr$ and L, characterize the Coulomb function,[5] the values of interest will be fairly large; η, for example, lies in the range 5−15. Tabulations in this particular range of parameters are either unavailable or incomplete[6] and the present work was undertaken to fill this need as far as feasible, with particular emphasis on large values of the parameter η. It extends and supplements the earlier work of Breit and his associates,[7] and of Abramowitz and Morse,[8] and in part runs parallel to or overlaps work of Newton,[9]

* Assisted by the Office of Ordnance Research, U. S. Army and by the joint program of the U. S. Office of Naval Research and U. S. Atomic Energy Commission.

† Now at the Rice Institute, Houston, Texas.

[1] Breit, Hull, and Gluckstern, Phys. Rev. **87**, 74 (1952); N. F. Ramsey, Phys. Rev. **83**, 659 (1951).

[2] L. D. Wyly and A. Zucker, Phys. Rev. **89**, 524 (1953).

[3] C. J. Mullin and E. Guth, Phys. Rev. **82**, 141 (1951); Ter-Martirosyan, J. Exp. Theoret. Phys. (USSR) **22**, 284 (1952); A. Bohr and B. R. Mottleson, Kgl. Danske Videnskab. Selskab, Mat.-fys. Medd. **27**, No. 16 (1953); K. Alder and A. Winther, Phys. Rev. **91**, 1578 (1953).

[4] T. Huus and C. Zupancic, Kgl. Danske Videnskab. Selskab, Mat.-fys. Medd. **28**, No. 1 (1953); C. McClelland and C. Goodman, Phys. Rev. **91**, 760 (1953); G. M. Temmer and N. P. Heydenburg, Phys. Rev. **94**, 1399 (1954); Sherr, Li, and Christy, Phys. Rev. **94**, 1076 (1954).

[5] Yost, Wheeler, and Breit, Phys. Rev. **49**, 174 (1936).

[6] The recent appearance of tables with $1\leq\eta\leq10$ [U. S. National Bureau of Standards Report No. 3033 (unpublished)] by C. E. Froberg and P. Rabinowitz is a welcome addition in this range.

[7] Yost, Wheeler, and Breit, reference 5; G. Breit and M. H. Hull, Jr., Phys. Rev. **80**, 392 (1950) and Phys. Rev. **80**, 561 (1950); Bloch, Hull, Broyles, Bouricius, Freeman, and Breit, Phys. Rev. **80**, 553 (1950).

[8] M. Abramowitz, *Tables of Coulomb Functions*, Vol. I, U. S. National Bureau of Standards Applied Mathematics Series, No. 17 (1952). Several expansions due to Mr. Abramowitz are discussed in the introduction, pp. xv–xxvii, and one due to P. M. Morse.

[9] T. D. Newton, Atomic Energy of Canada, Limited, Report CRT-526, 1952 (unpublished).

Tyson,[10] Feshbach, Shapiro, and Weisskopf,[11] and more recent work of Abramowitz.[12] Two different needs have been kept in mind in this work: The desirability of having formulas for calculating the functions at given points to high accuracy for starting numerical integrations, and of being able to obtain fairly good values of the functions for *any* value of ρ, η, L for estimates, checks, etc. The first expressed need is met here in series which converge rapidly for large η and for $L<\eta$, $L\sim\eta$, $L>\eta$ at two special values of ρ: $\rho=2\eta$ and $\rho=\rho_L\equiv\eta+[\eta^2+L(L+1)]^{\frac{1}{2}}$. These will be recognized, respectively, as the classical turning point for $L=0$ and for any L. The methods of obtaining the expansions start from the integral representations of the functions (given, in the form used here, by Bloch, Hull, Broyles, Bouricius, Freeman, and Breit[13]). Use of the special points indicated leads to simplifications in the work. Expansions for arbitrary values of ρ in the vicinity of the turning points are also obtained. A noteworthy lack of dependence on L is found, and for general values of $\rho<\rho_L$ extending down to fairly small fractions of ρ_L, it has been found possible to group the functions rather close to a "universal" curve from which values for any η, L can be read.

The second need is met by an approximation by means of Bessel functions of order one-third which is an extension of the procedure given by Morse and Feshbach.[14] Again the integral representation, or expressions obtained from it, is used to evaluate constants in the approximation. For the present case, it is found that the difference between the potential appearing in the equation actually solved by the usual Morse-Feshbach approximation and the true Coulomb potential is nearly proportional to ρ^{-2}. An improved approximation is easily obtained, therefore, by defining an effective value of L in the Morse-Feshbach potential. This approach is pursued, and comparisons of accuracy are made for several special values of ρ. The approximation suffers, of course, because its accuracy is not definitely assignable for arbitrary ρ, but its virtue is that it yields values of the function for any ρ with an error not expected to exceed one or two percent.

II. EXPANSIONS AT THE TURNING POINTS

Expansions at the turning points for $F_L(\rho)$ and $G_L(\rho)$ and their derivatives, valid for large values of η and moderate values of $L<\eta$, have been obtained from the integral representations for these functions given in the paper of Bloch, Hull, Broyles, Bouricius, Freeman, and Breit[13]—their equation (9). In what follows we shall use the notation of this paper. It is convenient to introduce a notation for the classical turning point, ρ_L, which has the value:

$$\rho_L=\eta+[\eta^2+L(L+1)]^{\frac{1}{2}}.$$

Letting $Y_L(\rho)=F_L(\rho)+iG_L(\rho)$, one finds from reference 13, Eq. (9), that:

$$\begin{aligned} Y_L(\rho)&=ie^{-i\rho}[(2L+1)!C_L\rho^L]^{-1}\\ &\quad\times\int_0^\infty t^{-i\eta+L}(t+2i\rho)^{i\eta+L}e^{-t}dt,\\ C_L&\equiv[2^L/(2L+1)!]\{[L^2+\eta^2][(L-1)^2+\eta^2]\cdots\\ &\quad\times[1+\eta^2]\}^{\frac{1}{2}}[2\pi\eta/(e^{2\pi\eta}-1)]^{\frac{1}{2}}. \end{aligned}\tag{1}$$

Consider first the expansion for $\rho=\rho_0=2\eta$. With $\rho=2\eta$ and the new variable z introduced by $t=i\rho_0(z-1)$, Eq. (1) becomes

$$Y_L(\rho_0)=-\frac{e^{-\pi\eta}(2\eta)^{L+1}}{[C_L(2L+1)!]}\int_1^{1-i\infty}(1-z^2)^L\times\exp\left\{i\eta\left[\ln\left(\frac{1+z}{1-z}\right)-2z\right]\right\}dz.\tag{2}$$

Introducing another change of variable defined by

$$2w^3/3=\ln[(1+z)/(1-z)]-2z,\tag{3}$$

one can put Eq. (2) in the form

$$Y_L(\rho_0)=\frac{e^{-\pi\eta}(2\eta)^{L+1}}{[(2L+1)!C_L]}\int_\Gamma[1-z^2(w)]^L\times\exp[\tfrac{2}{3}i\eta w^3]\frac{dz(w)}{dw}dw,\tag{4}$$

where the contour Γ is taken from $-i\infty$ to 0 and from 0 to ∞. The desired expansion now results when $[1-z^2(w)]^L dz(w)dw$ is expanded in a power series in w, and term-by-term integration is carried out. Upon taking the real and imaginary parts separately, the desired expansions for F_L and G_L are obtained. This procedure is quite straightforward and will not be discussed in detail. The series for G_L differs from that for F_L only in that some of the terms differ in sign and there is an over-all difference in size by the factor $\sqrt{3}$. Utilizing this fact, one can express the final results in the compact form

$$\begin{aligned} \begin{Bmatrix} F_L(\rho_0)\\ G_L(\rho_0)/\sqrt{3}\end{Bmatrix}&=\tfrac{1}{2}\Gamma(\tfrac{1}{3})\pi^{-\frac{1}{2}}(\tfrac{2}{3}\eta)^{1/6}(1-e^{-2\pi\eta})^{\frac{1}{2}}\\ &\times\{1\mp[3/35+L(L+1)/2]a\eta^{-4/3}-[2/225\\ &+L(L+1)/20]\eta^{-2}\mp[724/170\,625\\ &+L(L+1)/6300-L^2(L+1)^2/8]a\eta^{-10/3}+\cdots\}, \end{aligned}\tag{5}$$

[10] J. K. Tyson, thesis, Massachusetts Institute of Technology (1948) (unpublished).

[11] Feshbach, Shapiro, and Weisskopf, Atomic Energy Commission Reports NYO 3077, NDA 15B-5, 1953 (unpublished).

[12] M. Abramowitz and H. A. Antosiewicz, U. S. National Bureau of Standards Report No. 3225, 1954 (unpublished); M. Abramowitz and P. Rabinowitz (to be published).

[13] Bloch, Hull, Broyles, Bouricius, Freeman, and Breit, 4th reference of footnote 7.

[14] P. M. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw-Hill Book Company, Inc., New York, 1953), Vol. II, Chap. 9.

where

$$a\equiv(\tfrac{2}{3})^{\frac{1}{3}}\Gamma(\tfrac{2}{3})/\Gamma(\tfrac{1}{3}),$$

and where the upper sign before an expression on the right is to be taken with the upper function on the left, and similarly with the lower. This convention is utilized throughout the paper.

It should be mentioned that in obtaining Eq. (5) an expansion of the coefficient C_L [given in Eq. (1)] has been made. In addition, it might be of interest to note that the coefficients for the $\eta^{-2\nu/3}$ terms in the brackets with $\nu=6n+1$ and $\nu=6n+4$ always vanish.

Similar series for the derivatives, at $\rho=\rho_0$, can be obtained in two ways, either from the integral representation, Eq. (1), after differentiating, or from the recurrence relation, Eq. (11.5) of reference 13. The results are identical, of course, and have the form

$$\left\{\begin{matrix}F_L'(\rho_0)\\ G_L'(\rho_0)/\sqrt{3}\end{matrix}\right\}=\tfrac{1}{2}\Gamma(\tfrac{2}{3})\pi^{-\frac{1}{2}}(2\eta/3)^{-1/6}(1-e^{-2\pi\eta})^{-\frac{1}{2}}$$
$$\times\{\pm 1+(1/10a)\eta^{-\frac{1}{3}}\pm(1/3150)\eta^{-2}$$
$$+([359/173\,250-L^2(L+1)^2/16]/a)\eta^{-8/3}+\cdots\}. \quad (6)$$

The coefficients for the $\eta^{-2\nu/3}$ terms in brackets with $\nu=6n+2$ and $\nu=6n+5$ always vanish, so that the first omitted term in Eq. (6) is of order η^{-4} compared to unity.

Instead of evaluating all the functions at the same point, say ρ_0 as above, it is useful to evaluate the functions for a given L at their own turning point, ρ_L. It develops that the results one then obtains are sensibly independent of the value of L.

To accomplish this one first of all changes the variable in Eq. (1) by the substitution $t\equiv i\rho(z-1)$. Then $Y_L(\rho)$ becomes

$$Y_L(\rho)=[e^{-\pi\eta}\rho^{L+1}/C_L(2L+1)!]\int_{1-i\infty}^{1}e^{\varphi(z)}dz,$$
$$\varphi(z)\equiv L\ln(1-z^2)+i\eta\ln[(1+z)/(1-z)]-i\rho z. \quad (7)$$

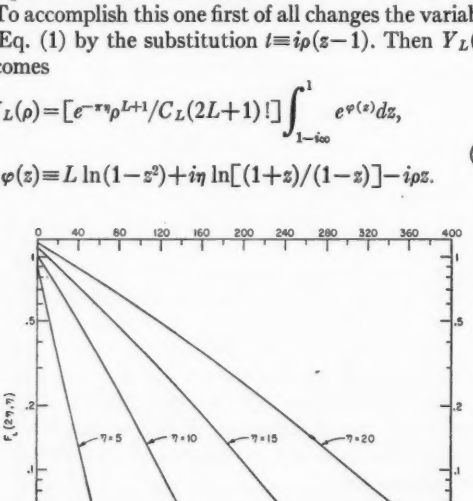


FIG. 1. $F_L(\rho_0)$ plotted as a function of L^2 for $\eta=5$, 10, 15, 20 on a semi-log scale. Equation (5) was used in the calculations for $L\lesssim\eta$, and Eq. (A3) when $L\gtrsim\eta$. The smooth variation of the curves allows easy interpolation in L for a given value of η.

If φ is differentiated with respect to z, one finds

$$d\varphi/dz=(1-z^2)^{-1}[i\rho(z+iL/\rho)^2+i(2\eta-\rho+L^2/\rho)]. \quad (8)$$

This suggests, first, that the variable be changed yet again from z to $s\equiv z+iL/\rho$ and, second, that one choose ρ to be that value for which $2\eta-\rho+L^2/\rho=0$. This is *not* the usual "classical" turning point value for ρ [in the sense that $L(L+1)\to L^2$ it is an even more "classical" value!] but the difference is rather slight. Designating this value of ρ by $\bar{\rho}_L$, one has $\bar{\rho}_L=\eta+[\eta^2+L^2]^{\frac{1}{2}}$. Although the work is considerably simpler for $\bar{\rho}_L$ than for ρ_L, the resulting formulas for F_L, G_L, and their derivatives assume the more compact form for $\rho=\rho_L$. The shift from $\bar{\rho}_L$ to ρ_L is easily accomplished by a Taylor's series. This apparently roundabout procedure outlined above is easier algebraically than the direct expansion at ρ_L.

Proceeding in this way, one finds that for $\rho=\bar{\rho}_L$

$$\varphi=\varphi_0+[(i\bar{\rho}_L s^3/3(1+L^2/\bar{\rho}_L^2)]+\cdots,$$
where
$$\varphi_0=L\ln(1+L^2/\bar{\rho}_L^2)-L+2\eta\tan^{-1}(L/\bar{\rho}_L). \quad (9)$$

In exact analogy to the methods used for the case where $\rho=2\eta$, one now makes still another change of variable defined by

$$\varphi\equiv\varphi_0+i\bar{\rho}_L w^3/[3(1+L^2/\bar{\rho}_L^2)].$$

The final result is

$$Y_L(\bar{\rho}_L)=\frac{e^{(\varphi_0-\pi\eta)}\bar{\rho}_L{}^{L+1}}{(2L+1)!C_L}$$
$$\times\int_\Gamma\exp\left(\frac{i\rho w^3}{3(1+L^2/\bar{\rho}_L^2)}\right)\left(\frac{dz}{dw}\right)dw. \quad (10)$$

The contour Γ is the same as for Eq. (4), and the details of the work are exactly as earlier. The results are

$$\left\{\begin{matrix}F_L(\bar{\rho}_L)\\ G_L(\bar{\rho}_L)/\sqrt{3}\end{matrix}\right\}=\tfrac{1}{2}\Gamma(\tfrac{1}{3})\pi^{-\frac{1}{2}}(\bar{\rho}_L/3)^{1/6}(1+L^2/\bar{\rho}_L^2)^{-1/6}$$
$$\times\left\{1\mp\bar{\rho}_L{}^{-\frac{1}{3}}b(1+L^2/\bar{\rho}_L^2)^{-\frac{1}{3}}(L/\bar{\rho}_L)\right.$$
$$+\frac{\bar{\rho}_L{}^{-1}(L/\bar{\rho}_L)(2L^2/\bar{\rho}_L^2-1)}{10(1+L^2/\bar{\rho}_L^2)^2}$$
$$\mp\bar{\rho}_L{}^{-4/3}b(b/35-11L^2/35\bar{\rho}_L^2$$
$$\left.+L^4/140\bar{\rho}_L^4)+\cdots\right\}, \quad (11)$$

$$\left\{\begin{matrix}F_L'(\bar{\rho}_L)\\ G_L'(\bar{\rho}_L)/\sqrt{3}\end{matrix}\right\}=\tfrac{1}{2}\Gamma(\tfrac{2}{3})\pi^{-\frac{1}{2}}(\bar{\rho}_L/3)^{-1/6}(1+L^2/\rho_L^2)^{1/6}$$
$$\times\left\{\pm 1+\frac{\bar{\rho}_L{}^{-\frac{1}{3}}(2-L^2/\bar{\rho}_L^2)}{10b(1+L^2/\bar{\rho}_L^2)^{4/3}}+\cdots\right\}, \quad (12)$$

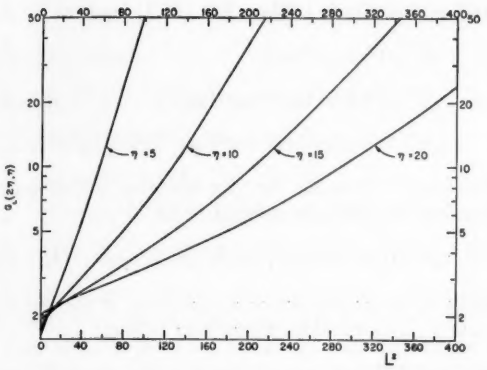


FIG. 2. $G_L(\rho_0)$ plotted as a function of L^2 for $\eta=5$, 10, 15, 20 on a semilog scale. Equation (5) was used in the calculations for $L\lesssim\eta$, and Eq. (A4) when $L\gtrsim\eta$. The smooth variation allows easy interpolation in L for a given value of η.

where $b=3^{\frac{1}{3}}\Gamma(\frac{2}{3})/\Gamma(\frac{1}{3})$. The next expected term in Eq. (11) and the next two in Eq. (12) contain zero factors.

It is now quite simple to obtain the desired expansions at the turning points by using a Taylor series for the shift from $\bar{\rho}_L$ to ρ_L. The results are[15]

$$\begin{Bmatrix} F_L(\rho_L) \\ G_L(\rho_L)/\sqrt{3} \end{Bmatrix} = \tfrac{1}{2}\Gamma(\tfrac{1}{3})\pi^{-\frac{1}{2}}(\rho_L/3)^{1/6}[1+L(L+1)/\rho_L^2]^{-1/6}$$

$$\times\{1\mp\rho_L^{-4/3}(6b/35)[1+L(L+1)/\rho_L^2]^{-8/3}$$

$$\times[1+4L(L+1)/\rho_L^2+3L^2(L+1)^2/2\rho_L^4]+\cdots\}, \quad (13)$$

$$\begin{Bmatrix} F_L'(\rho_L) \\ G_L'(\rho_L)/\sqrt{3} \end{Bmatrix} = \tfrac{1}{2}\Gamma(\tfrac{2}{3})\pi^{-\frac{1}{2}}(\rho_L/3)^{-1/6}[1+L(L+1)/\rho_L^2]^{1/6}$$

$$\times\left\{\pm 1+\frac{\rho_L^{-\frac{1}{3}}[1+2L(L+1)/\rho_L^2]}{5b[1+L(L+1)/\rho_L^2]^{4/3}}+\cdots\right\}. \quad (14)$$

The dependence of the above results on L can be put in evidence by expanding the various factors containing L in powers of L/η, which will be presumed small. Thus,

$$\{\rho_L/[1+L(L+1)]\rho_L^2]\}^{\pm 1/6} = (2\eta)^{\pm 1/6}[1\pm L^2(L+1)^2/96\eta^4+\cdots].$$

Introducing this into Eqs. (13) and (14), one finds that $F_L(\rho_L)$ and $G_L(\rho_L)$ are independent of L to about 1 percent even when $L=\eta$, for η large.

Unlike Eqs. (5) and (6), which required $L<\eta$, Eqs. (13) and (14) [and Eqs. (11) and (12), as well] are valid for arbitrary values of L, and, in fact, can be used to obtain the asymptotic forms for $L\to\infty$. In this limit one finds

$$\begin{Bmatrix} F_L(\rho_L) \\ G_L(\rho_L)/\sqrt{3} \end{Bmatrix} \to \tfrac{1}{2}\Gamma(\tfrac{1}{3})\pi^{-\frac{1}{2}}(L/6)^{1/6},$$

$$\begin{Bmatrix} F_L'(\rho_L) \\ G_L'(\rho_L)/\sqrt{3} \end{Bmatrix} \to \pm\tfrac{1}{2}\Gamma(\tfrac{2}{3})\pi^{-\frac{1}{2}}(L/6)^{-1/6},$$

with $L\to\infty$.

When L and η are of the same order, and large compared to unity, ρ_L is about 25 percent larger than ρ_0 and the difference between ρ_L and ρ_0 increases as L becomes greater than η. As a consequence the regular and irregular properties of F_L and G_L, respectively, begin to appear markedly at ρ_0, so that they are no longer of the same order numerically. This is, of course, the basis for the requirement that L be moderately small compared to η for Eqs. (5) and (6) to be valid. If one desires expansions at ρ_0 for cases when this restriction does not obtain, it is clear that a very much different approach is required. It is convenient in this case then to study F_L and G_L from the standpoint of their own integral representations. This case, where $L\gtrsim\eta$, will not be discussed here; for reference, however, some results appropriate to this region are collected in the Appendix.

Since the calculations of Barfield and Broyles,[16] for $F_0(\rho_0)$, $F_0'(\rho_0)$, and $G_L(\rho_0)$, led initially to the work described above, one of the first applications was to verify their results. It was found that for $\eta\gtrsim 10$ the first two nonvanishing terms of Eqs. (5), (6) for $L=0$ gives the Barfield-Broyles values to the accuracy written by them, and for $\eta>30$ the first term is sufficient. This result agrees with Newton,[9] who made the same comparison.

Calculations have also been performed for $\eta\gtrsim 5$ and arbitrary L. The convergence of Eqs. (5) and (6) rapidly becomes poorer as L increases, and for high L the alternative results given in the Appendix were necessary.

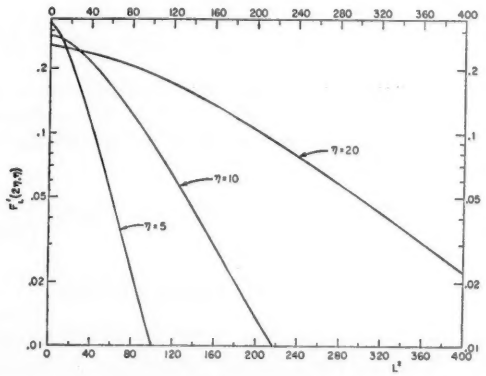


FIG. 3. $F_L'(\rho_0)$ plotted as a function of L^2 for $\eta=5$, 10, 20 on a semilog scale. Equation (6) was used in the calculations for $L\lesssim\eta$, and Eq. (A5) when $L\gtrsim\eta$. The smooth variation of the curves allows easy interpolation in L for a given value of η.

[15] M. Abramowitz and P. Rabinowitz, second reference of footnote 12, have obtained, in a preprint recently received by us, an expansion at ρ_0 for $L=0$ to which our Eqs. (13) and (14) reduce in that special case.

[16] W. D. Barfield and A. A. Broyles, Phys. Rev. 88, 892 (1954).

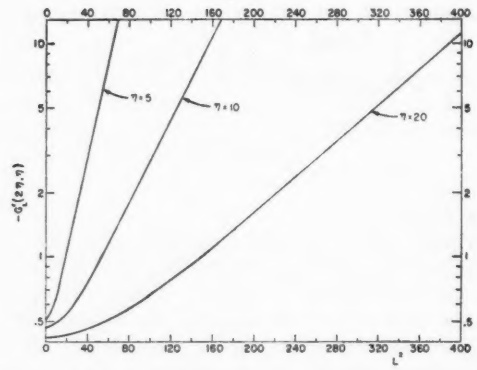


FIG. 4. $-G_L'(\rho_0)$ plotted as a function of L^2 for $\eta=5$, 10, 20 on a semilog scale. Equation (6) was used in the calculations for $L\leqslant\eta$, and Eq. (A5) when $L\geqslant\eta$. The smooth variation of the curves allows easy interpolation in L for a given value of η.

The results of this numerical work are displayed in Figs. 1–4.

Further calculations have been performed for $F_L(\rho_L)$, $G_L(\rho_L)$ and their derivatives by using Eqs. (13) and (14). The range of validity of these equations is considerable, since they hold for arbitrary values of L/η provided only that η is large. In the parameter range $\eta\sim10$, the first term alone in Eq. (13) is sufficient to give $F_L(\rho_L)$ or $G_L(\rho_L)$ to 2 percent, with accuracy improving as L increases. The results of the calculations are plotted in Figs. 5 and 6. It will be noted that these figures demonstrate the lack of sensitivity to the value of L remarked on earlier.

III. EXPANSIONS VALID IN THE VICINITY OF THE TURNING POINT

In the vicinity of the turning point, ρ_L, an expansion can be obtained by successive approximation with the aid of the Green's function for the operator d^2/dx^2+x. The equation satisfied by F_L and G_L is

$$d^2Y_L/d\rho^2+[1-2\eta/\rho-L(L+1)/\rho^2]Y_L=0. \tag{15}$$

It is convenient to change the variable from ρ to $x=(\rho-\rho_L)[1/\rho_L+L(L+1)/\rho_L^2]^{\frac{1}{3}}$. Then the coefficient of Y_L in Eq. (15) can be expanded as a power series in x, so that Eq. (15) becomes

$$\begin{aligned} &d^2Y_L/dx^2+x(1+\lambda_1\rho_L^{-\frac{1}{3}}x+\lambda_2\rho_L^{-4/3}x^2+\cdots)Y_L=0, \\ &\lambda_n=(-1)^n[1+L(L+1)/\rho_L^2]^{-(n+2)/3} \\ &\qquad\times[1+nL(L+1)/\rho_L^2]. \end{aligned} \tag{15.1}$$

Two independent solutions of Eq. (15.1) will be constructed having the following behavior at $x=0$: $f(x)$ shall have zero value and unit slope at $x=0$, $g(x)$ shall have unit value and zero slope at $x=0$ (slope here means derivative with respect to x). Expand $f(x)$ as

$$f(x)=f_0(x)+\rho_L^{-\frac{1}{3}}f_1(x)+\cdots,$$

[and $g(x)$ similarly]. Then Eq. (15.1) leads to

$$\begin{aligned} &f_0''(x)+xf_0(x)=0, \\ &f_1''(x)+xf_1(x)=-\lambda_1x^2f_0\rho_L^{-\frac{1}{3}}, \\ &f_2''(x)+xf_2(x)=-\lambda_1x^2f_1\rho_L^{-\frac{1}{3}}-\lambda_2x^3f_0\rho_L^{-4/3}. \end{aligned} \tag{16}$$

The Green's function for the operator (d^2/dx^2+x) is therefore required. It is found to be

$$G(x,x')=(2\pi/3^{\frac{1}{2}})x_<^{\frac{1}{2}}J_{\frac{1}{3}}(2x_<^{\frac{3}{2}}/3)x_>^{\frac{1}{2}}J_{-\frac{1}{3}}(2x_>^{\frac{3}{2}}/3), \tag{17}$$

where $x_<$ is the smaller of x and x', $x_>$ is the larger of x and x'. The solution proceeds as follows:

$$f_0(x)=x^{\frac{1}{2}}J_{\frac{1}{3}}(\tfrac{2}{3}x^{\frac{3}{2}}),$$

$$f_1(x)=-\lambda_1\int f_0(x')x'^{\frac{3}{2}}G(x,x')dx', \tag{18}$$

$$\begin{aligned} f_2(x)=&-\lambda_1\int f_1(x')x'^{\frac{3}{2}}G(x,x')dx' \\ &-\lambda_2\int f_0(x')x'^3G(x,x')dx', \end{aligned}$$

which leads to integrals of the forms

$$\int t^{(3n+5)/3}J_{\pm\frac{1}{3}}(t)J_{\pm\frac{1}{3}}(t)dt.$$

These indefinite integrals are special cases of the Lommel integrals in Sec. 5.12 of Watson's *Bessel Functions*, or may be evaluated with their help after application of a reduction formula due to Schafheitlin (Sec. 5.14 of Watson) and extensions obtained in a manner analogous to that of Schafheitlin. The resulting expressions are sums of products of 3 Bessel functions, but have a common factor which arranges itself into a Wronskian relation, leaving sums of single Bessel functions. Suitable portions of zero order solutions had to be added to ensure the proper behavior of the high order solutions at $x=0$, since this procedure was found to be easier than adjusting the limits in Eq. (18) properly.

In the case of $f(x)$, for example, $f_0(x)$ has the desired behavior at $x=0$, and $f_1(x)$, $f_2(x)$ were adjusted so as not to interfere with this.

The results of these calculations are:

$$\begin{aligned} &f_0(x)=x^{\frac{1}{2}}J_{\frac{1}{3}}(\tfrac{2}{3}x^{\frac{3}{2}}), \\ &f_1(x)=-(\lambda_1/5)x^3J_{4/3}(\tfrac{2}{3}x^{\frac{3}{2}}), \\ &f_2(x)=(\lambda_1^2/350)[(30x^4-90x)J_{-\frac{1}{3}}(\tfrac{2}{3}x^{\frac{3}{2}}) \\ &\qquad+(7x^{11/2}-45x^{5/2})J_{\frac{1}{3}}(\tfrac{2}{3}x^{\frac{3}{2}})+90bx^{\frac{1}{2}}J_{-\frac{1}{3}}(\tfrac{2}{3}x^{\frac{3}{2}})] \\ &\qquad+(\lambda_2/14)[(2x^4-6x)J_{-\frac{1}{3}}(\tfrac{2}{3}x^{\frac{3}{2}}) \\ &\qquad\qquad-3x^{5/2}J_{\frac{1}{3}}(\tfrac{2}{3}x^{\frac{3}{2}})]+6bx^{\frac{1}{2}}J_{-\frac{1}{3}}(\tfrac{2}{3}x^{\frac{3}{2}})], \end{aligned} \tag{19}$$

where

$$b=3^{\frac{1}{2}}\Gamma(\tfrac{2}{3})/\Gamma(\tfrac{1}{3}),$$

and

$$g_0(x)=x^{\frac{1}{2}}J_{-\frac{1}{3}}(\tfrac{2}{3}x^{\frac{3}{2}}),$$

$$g_1(x)=(\lambda_1/5)[x^3J_{-4/3}(\tfrac{2}{3}x^{\frac{3}{2}})+(1/b)x^{\frac{1}{2}}J_{\frac{1}{3}}(\tfrac{2}{3}x^{\frac{3}{2}})],$$

$$\begin{aligned}g_2(x)=(\lambda_1^2/350)&[(30x^4-90x)J_{\frac{1}{3}}(\tfrac{2}{3}x^{\frac{3}{2}})\\&-(7x^{11/2}-45x^{5/2})J_{-\frac{1}{3}}(\tfrac{2}{3}x^{\frac{3}{2}})]\\&-(\lambda_1^2/25b)x^3J_{4/3}(\tfrac{2}{3}x^{\frac{3}{2}})-(\lambda_2/14)\\&\times[(2x^4-6x)J_{\frac{1}{3}}(2x^{\frac{3}{2}})+3x^{5/2}J_{-\frac{1}{3}}(2x^{\frac{3}{2}})].\end{aligned}\quad(20)$$

The expressions in Eqs. (19) and (20) are convenient for positive values of x, i.e., for $\rho>\rho_L$.

Equivalent expressions for $\rho<\rho_L$ may be written in terms of the modified Bessel functions of the first kind and order: $\pm\frac{1}{3}$, $\pm\frac{2}{3}$, $\pm$ etc. These expressions are:

$$f_0=-y^{\frac{1}{2}}I_{\frac{1}{3}}(\tfrac{2}{3}y^{\frac{3}{2}}),$$

$$f_1=(\lambda_1/5)y^3I_{4/3}(\tfrac{2}{3}y^{\frac{3}{2}}),$$

$$\begin{aligned}f_2=(\lambda_1^2/350)&[(30y^4+90y)I_{-\frac{1}{3}}(\tfrac{2}{3}y^{\frac{3}{2}})\\&-(7y^{11/2}+45y^{5/2})I_{\frac{1}{3}}(\tfrac{2}{3}y^{\frac{3}{2}})-90by^{\frac{1}{2}}I_{-\frac{1}{3}}(\tfrac{2}{3}y^{\frac{3}{2}})]\\&+(\lambda_2/14)[-(2y^4+6y)I_{-\frac{1}{3}}(\tfrac{2}{3}y^{\frac{3}{2}})\\&+3y^{5/2}I_{\frac{1}{3}}(\tfrac{2}{3}y^{\frac{3}{2}})+6by^{\frac{1}{2}}I_{-\frac{1}{3}}(\tfrac{2}{3}y^{\frac{3}{2}})],\end{aligned}\quad(21)$$

where

$$y=(\rho_L-\rho)[1/\rho_L+L(L+1)/\rho_L^3]^{\frac{1}{3}},$$

and

$$g_0=y^{\frac{1}{2}}I_{-\frac{1}{3}}(\tfrac{2}{3}y^{\frac{3}{2}}),$$

$$g_1=-(\lambda_1/5)[y^3I_{-4/3}(\tfrac{2}{3}y^{\frac{3}{2}})+(1/b)y^{\frac{1}{2}}I_{\frac{1}{3}}(\tfrac{2}{3}y^{\frac{3}{2}})],$$

$$\begin{aligned}g_2=(\lambda_1^2/350)&[-(30y^4+90y)I_{\frac{1}{3}}(\tfrac{2}{3}y^{\frac{3}{2}})\\&+(7y^{11/2}+45y^{5/2})I_{-\frac{1}{3}}(\tfrac{2}{3}y^{\frac{3}{2}})]+(\lambda_2/25b)y^3I_{4/3}\\&\times(\tfrac{2}{3}y^{\frac{3}{2}})+(b/14)[(2y^4+6y)I_{\frac{1}{3}}(\tfrac{2}{3}y^{\frac{3}{2}})\\&-3y^{5/2}I_{-\frac{1}{3}}(\tfrac{2}{3}y^{\frac{3}{2}})].\end{aligned}\quad(22)$$

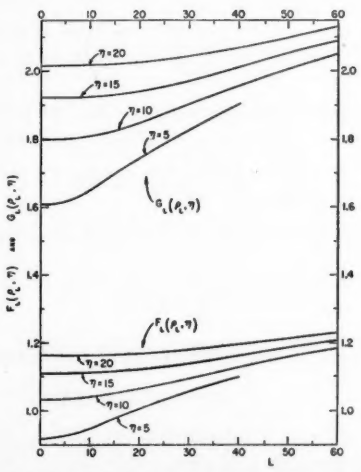


FIG. 5. $F_L(\rho_L)$ and $G_L(\rho_L)$ plotted as a function of L for $\eta=5$, 10, 15, 20. Equation (13) was used in the calculations for all values of L. The slow variation of the functions with L is illustrated.

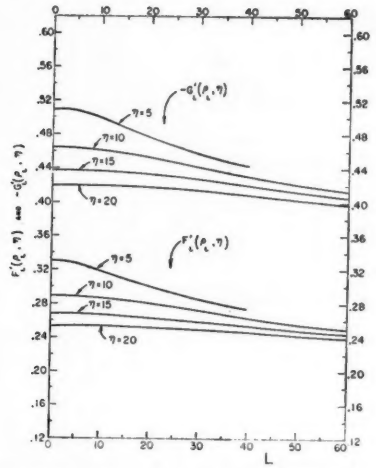


FIG. 6. $F_L'(\rho_L)$ and $-G_L'(\rho_L)$ plotted as a function of L for $\eta=5$, 10, 15, 20. Equation (14) was used in the calculations for all values of L. The slow variation of the functions with L is illustrated.

The desired formulas for F_L and G_L are then

$$\begin{aligned}F_L(\rho)&=F_L(\rho_L)g(x)\\&\quad+F_L'(\rho_L)[1/\rho_L+L(L+1)/\rho_L^3]^{\frac{1}{3}}f(x),\\G_L(\rho)&=G_L(\rho_L)g(x)\\&\quad+G_L'(\rho_L)[1/\rho_L+L(L+1)/\rho_L^3]^{\frac{1}{3}}f(x).\end{aligned}\quad(23)$$

The forms of $F_L(\rho_L)$, $G_L(\rho_L)$, etc. of Eqs. (13) and (14) may be used with Eqs. (23) to write explicit expansions of $F_L(\rho)$, $G_L(\rho)$. These are

$$\begin{aligned}\begin{Bmatrix}F_L(\rho)\\G_L(\rho)/\sqrt{3}\end{Bmatrix}&=\frac{\pi^{\frac{1}{2}}}{3}\left(\frac{\rho_L}{1+L(L+1)/\rho_L^2}\right)^{1/6}\\&\times\Big\{x^{\frac{1}{2}}[J_{-\frac{1}{3}}(\tfrac{2}{3}x^{\frac{3}{2}})\pm J_{\frac{1}{3}}(\tfrac{2}{3}x^{\frac{3}{2}})]\\&-\rho_L^{-1}\left(\frac{1+2L(L+1)/\rho_L^2}{5[1+L(L+1)/\rho_L^2]^{4/3}}\right)x^3\\&\times[J_{-4/3}(\tfrac{2}{3}x^{\frac{3}{2}})\mp J_{4/3}(\tfrac{2}{3}x^{\frac{3}{2}})]\\&+\rho_L^{-4/3}\frac{[1+2L(L+1)/\rho_L^2]^2}{[1+L(L+1)/\rho_L^2]^{8/3}}\Big[\left(\frac{3x^4-9x}{35}\right)\\&\times[J_{\frac{1}{3}}(\tfrac{2}{3}x^{\frac{3}{2}})\mp J_{-\frac{1}{3}}(\tfrac{2}{3}x^{\frac{3}{2}})]+\left(-\frac{x^{11/2}}{50}+9\frac{x^{5/2}}{70}\right)\\&\times[J_{-\frac{1}{3}}(\tfrac{2}{3}x^{\frac{3}{2}})\pm J_{\frac{1}{3}}(\tfrac{2}{3}x^{\frac{3}{2}})]\Big]\\&+\rho_L^{-4/3}\left(\frac{1+3L(L+1)/\rho_L^2}{[1+L(L+1)/\rho_L^2]^{5/3}}\right)\Big[-\left(\frac{x^4-3x}{7}\right)\\&\times[J_{\frac{1}{3}}(\tfrac{2}{3}x^{\frac{3}{2}})\mp J_{-\frac{1}{3}}(\tfrac{2}{3}x^{\frac{3}{2}})]\\&\quad-\left(\frac{3x^{5/2}}{14}\right)[J_{-\frac{1}{3}}(\tfrac{2}{3}x^{\frac{3}{2}})\pm J_{\frac{1}{3}}(\tfrac{2}{3}x^{\frac{3}{2}})]\Big]+\cdots\Big\},\end{aligned}\quad(24)$$

for $\rho-\rho_L>0$, and

$$\left\{\begin{matrix}F_L(\rho)\\ G_L(\rho)/\sqrt{3}\end{matrix}\right\}=\frac{\pi^{\frac{1}{2}}}{3}\left(\frac{\rho_L}{1+L(L+1)/\rho_L{}^2}\right)^{1/6}$$
$$\times\Big\{y^{\frac{1}{2}}[I_{-\frac{1}{3}}(\tfrac{2}{3}y^{\frac{3}{2}})\mp I_{\frac{1}{3}}(\tfrac{2}{3}y^{\frac{3}{2}})]$$
$$+\rho_L{}^{-1}\left(\frac{1+2L(L+1)/\rho_L{}^2}{5[1+L(L+1)/\rho_L{}^2]^{4/3}}\right)y^3$$
$$\times[I_{-4/3}(\tfrac{2}{3}y^{\frac{3}{2}})\mp I_{4/3}(\tfrac{2}{3}y^{\frac{3}{2}})]$$
$$+\rho_L{}^{-4/3}\left(\frac{[1+2L(L+1)/\rho_L{}^2]^2}{[1+L(L+1)/\rho_L{}^2]^{8/3}}\right)\Big[-\left(\frac{3y^4-9y}{35}\right)$$
$$\times[I_{\frac{1}{3}}(\tfrac{2}{3}y^{\frac{3}{2}})\mp I_{-\frac{1}{3}}(\tfrac{2}{3}y^{\frac{3}{2}})]$$
$$+\left(\frac{y^{11/2}}{50}+\frac{9y^{5/2}}{70}\right)[I_{-\frac{1}{3}}(\tfrac{2}{3}y^{\frac{3}{2}})\mp I_{\frac{1}{3}}(\tfrac{2}{3}y^{\frac{3}{2}})]\Big]$$
$$+\rho_L{}^{-4/3}\left(\frac{1+3L(L+1)/\rho_L{}^2}{[1+L(L+1)/\rho_L{}^2]^{5/3}}\right)\Big[\left(\frac{y^4+3y}{7}\right)$$
$$\times[I_{\frac{1}{3}}(\tfrac{2}{3}y^{\frac{3}{2}})\mp I_{-\frac{1}{3}}(\tfrac{2}{3}y^{\frac{3}{2}})]$$
$$-\left(\frac{3y^{5/2}}{14}\right)[I_{-\frac{1}{3}}(\tfrac{2}{3}y^{\frac{3}{2}})\mp I_{\frac{1}{3}}(\tfrac{2}{3}y^{\frac{3}{2}})]\Big]+\cdots\Big\},\quad(25)$$

for $\rho-\rho_L<0$.

The results given in Eqs. (24) and (25) may also be obtained directly from the integral representation of Eq. (1), by suitably expanding the integrand. The leading term, for example, yields an integral that can be put in the form of Airy's integral and leads to the Bessel functions of order one-third.[17] It may be noted in Eq. (25) that the functions with positive and negative orders occur in the expression for $F_L(\rho)$ in the proper combination to make $K_\nu(\frac{2}{3}x^{\frac{3}{2}})$.

It has already been brought out, in connection with Eqs. (13) and (14), that the dependence on L of the functions $F_L(\rho_L)$ and $G_L(\rho_L)$ is in the fourth power of L/η, in particular, the dependence is $(1+L^2(L+1)^2/96\eta^4)$ for large η. Since

$$y=(\rho_L-\rho)[1/\rho_L+L(L+1)/\rho_L{}^3]^{\frac{1}{3}}$$
$$\cong[(\rho_L-\rho)/(2\eta)^{\frac{1}{3}}][1-L^2(L+1)^2/48\eta^4],$$

Eq. (25) leads one to expect that for a fair range of parameters, $F_L(\rho)$ and $G_L(\rho)$ could be represented by "universal" curves fitted respectively to $F_L(\rho)/\eta^{1/6}$ plotted against $(\rho_L-\rho)/\eta^{\frac{1}{3}}$ and to $G_L(\rho)/\eta^{1/6}$ plotted against the same variable. Figure 7 shows such a set of curves, and illustrates indeed that the functions bunch fairly closely when plotted as suggested by the limiting forms of Eq. (25). The functions for $L=10$, $\eta=5$ deviate most, as is expected since L/η is then 2. The bunching is expected to become yet more pronounced as η exceeds L by larger and larger factors. At $y=0$, the values of $F_L(\rho_L)/\eta^{1/6}$ and $G_L(\rho_L)/\eta^{1/6}$ exhibit the even more pronounced (L,η) independence indicated in the discussion of Eq. (13).

Expansions similar to Eqs. (24) and (25), but valid in the vicinity of ρ_0 rather than ρ_L, have been given by Newton.[9] His expansion converges for small values of L. For $L\sim\eta$ or larger, expansions useful near ρ_0 may be obtained by other means indicated in the Appendix.

[17] M. Abramowitz and H. A. Antosiewicz, first reference of footnote 12, have obtained an expansion of $F_0(\rho)$ and $G_0(\rho)$ in terms of the Airy integrals for $|\rho-\rho_0|<\rho_0$.

IV. APPROXIMATE WAVE FUNCTIONS

The discussion of the previous sections has had as its object the enumeration of convenient methods for determining the Coulomb wave functions for large η to an arbitrarily given accuracy in any one of several regions of interest. There exists, however, a quite different objective, namely the need for simple functional forms for the Coulomb wave functions valid to reasonable accuracy, say $\sim$1 percent, over extended regions for ρ. Although somewhat too crude for our purposes, the JWKB approximation is typical to this approach. An improved approximation along these lines has been given by Morse and Feshbach in terms of the Bessel functions of order one-third. It is characteristic of both these points of view that one seeks to relate the solution of the problem at hand to the known solutions of a differential equation that is approximately the same, the physical basis of the approximation being in all cases the small change in the potential over distances of the order of a wavelength. In order to obtain an approximate solution to

$$d^2y/dx^2+g(x)y=0,\quad(26)$$

it is convenient to make a transformation,

$$y=CuY\left(\int^x u^{-2}dx\right),\quad(27)$$

where C is an arbitrary constant. One has then

$$\frac{d^2y}{dx^2}=\left(\frac{1}{u}\frac{d^2u}{dx^2}+\frac{1}{u^4}\frac{Y''}{Y}\right)y.\quad(27.1)$$

So far $Y(z)$ is an arbitrary function of argument

$$z=\int^x u^{-2}dx.\quad(28)$$

Once chosen, Y determines a function χ through

$$Y''(z)+\chi(z)Y(z)=0,\quad(29)$$

and in terms of χ substitution of (29) in (27.1) and identification with (26) gives

$$g(x)=\frac{1}{u^4}\chi\left(\int^x\frac{dx}{u^2}\right)-\frac{1}{u}\frac{d^2u}{dx^2}.\quad(30)$$

Conversely, if $\chi(z)$ is an assigned functional form, then a suitably chosen u inserted in (30) gives a solution of Eq. (26) provided one solves (29) so as to be able to insert Y in (27). Even if $g(x)$ cannot be represented exactly by means of Eq. (30) there are available adjustments in both u and χ for its approximate representation. In terms of z

$$g(x)=(dz/dx)^2\chi(z(x))+\Delta, \qquad \Delta=-(dz/dx)^{\frac{1}{2}}\frac{d^2[(dz/dx)^{-\frac{1}{2}}]}{dx^2}, \tag{30.1}$$

and

$$y=(dz/dx)^{-\frac{1}{2}}Y(z(x)). \tag{30.2}$$

In the latter form z is seen to be a generalization of the phase of the JWKB approximation.

By specializing to $\chi=1$, $a=$const, Eq. (30) becomes

$$g=\frac{a}{u^4}-\frac{1}{u}\frac{d^2u}{dx^2}. \tag{31}$$

An approximate solution is then $u\cong(a/g)^{\frac{1}{4}}$ which gives

$$y\cong C(a/g)^{\frac{1}{4}}Y\left(\int^x (g/a)^{\frac{1}{2}}dx\right) =C'g^{-\frac{1}{4}}\sin\left(\int^x g^{\frac{1}{2}}dx+C''\right), \tag{31.1}$$

where C', C'' are arbitrary constants. If Eq. (31) is solved for u more accurately, correction terms to the one term JWKB formula are obtained.

The JWKB approximation is an immediate generalization of a representation of $g(x)$ by constant steps. An improvement in convergence of successive terms is obtained[18] if $g(x)$ is approximated by a set of straight line segments. Such an approximation is found from the transformation under discussion by setting

$$\chi(z)=z \tag{31.2}$$

so that according to (30.1)

$$g(x)=z(dz/dx)^2+\Delta. \tag{31.3}$$

Again if Δ is not too important,

$$z^{\frac{3}{2}}=\frac{3}{2}\int^x g^{\frac{1}{2}}dx, \tag{31.4}$$

and

$$Y(z)=z^{\frac{1}{2}}[AJ_{\frac{1}{3}}(\tfrac{2}{3}z^{\frac{3}{2}})+BJ_{-\frac{1}{3}}(\tfrac{2}{3}z^{\frac{3}{2}})], \tag{31.5}$$

so that according to (30.2),

$$y=(\varphi/g^{\frac{1}{2}})^{\frac{1}{4}}[A'J_{\frac{1}{3}}(\varphi)+B'J_{-\frac{1}{3}}(\varphi)], \tag{31.6}$$

where

$$\varphi=\int^x g^{\frac{1}{2}}dx. \tag{31.7}$$

[18] R. E. Langer, Phys. Rev. **51**, 669 (1937).

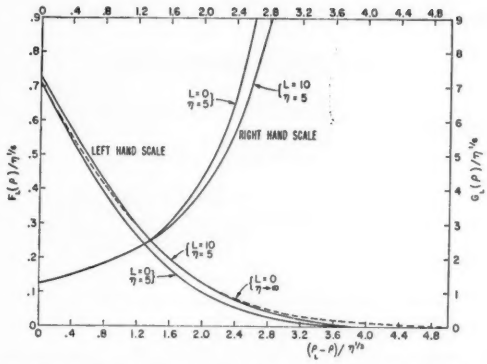


FIG. 7. $F_L(\rho)/\eta^{1/6}$ and $G_L(\rho)/\eta^{1/6}$ plotted as a function of $(\rho_L-\rho)/\eta^{1/3}$. These curves illustrate the possibility of representing the functions by "universal" curves, as suggested by Eq. (25). This tendency to group is lost as L exceeds η more and more and is emphasized as η exceeds L. The calculations were done by numerical integrations started at $\rho=\rho_L$ using Eqs. (13) and (14).

It should be noted that whenever the approximation is made by setting $\Delta=0$, then according to (30.1)

$$\varphi=\int^x g^{\frac{1}{2}}dx\cong\int\chi^{\frac{1}{2}}dz, \tag{32}$$

so that the geometrical optics approximation of the solutions to (26) and (29) contain the same phase. The approximation of Eq. (31.6) has been first obtained by Morse and Feshbach[14] who have also worked out the differential equation satisfied by the right side of (31.6). They have found in this special case a relation equivalent to Eq. (30.1) above.

Other approximations applicable in special situations are obtainable by the same method. If, for example, $g(x)$ can be approximately represented by a power of x then it is useful to take $\chi(z)=z^\nu$. In this case

$$y\cong(\varphi/g^{\frac{1}{2}})^{\frac{1}{2}}[AJ_p(\varphi)+BJ_{-p}(\varphi)], \qquad p=1/(\nu+2). \tag{32.1}$$

For the Coulomb wave functions, the turning point has a zero of first order in g and consequently the Morse-Feshbach approximation is indicated. When the constants are adjusted to fit the boundary conditions and the values at the turning point, one finds:

For $\rho\gtrsim\eta+[\eta^2+L(L+1)]^{\frac{1}{2}}$,

$$g^{\frac{1}{2}}=[1-2\eta/\rho-L(L+1)/\rho^2]^{\frac{1}{2}},$$

$$\varphi=\rho g^{\frac{1}{2}}-\eta\ln\left(\frac{(\rho-\eta+\rho g^{\frac{1}{2}})}{[\eta^2+L(L+1)]^{\frac{1}{2}}}\right) -[L(L+1)]^{\frac{1}{2}}\sin^{-1}\left(\frac{[L(L+1)g]^{\frac{1}{2}}}{[\eta^2+L(L+1)]^{\frac{1}{2}}}\right),$$

$$\left\{\begin{matrix}F_L(\rho)\\ G_L(\rho)/\sqrt{3}\end{matrix}\right\}\cong(\pi\varphi/6g^{\frac{1}{2}})^{\frac{1}{2}}[J_{-\frac{1}{3}}(\varphi)\pm J_{\frac{1}{3}}(\varphi)]. \tag{33}$$

For $\rho \leqslant \eta+[\eta^2+L(L+1)]^{\frac{1}{2}}$,

$$g^{\frac{1}{2}}=[L(L+1)/\rho^2+2\eta/\rho-1]^{\frac{1}{2}},$$

$$\varphi=i\left\{\rho g^{\frac{1}{2}}-\eta\tan^{-1}\left(\frac{\rho g^{\frac{1}{2}}}{(\rho-\eta)}\right)\right.$$

$$\left.-[L(L+1)]^{\frac{1}{2}}\sinh^{-1}\left(\frac{[L(L+1)g]^{\frac{1}{2}}}{[\eta^2+L(L+1)]^{\frac{1}{2}}}\right)\right\} \quad (33.1)$$

for $\rho>\eta$, and

$$=i\left\{\rho g^{\frac{1}{2}}-\pi\eta-\eta\tan^{-1}\left(\frac{\rho g^{\frac{1}{2}}}{\rho-\eta}\right)\right.$$

$$\left.-[L(L+1)]^{\frac{1}{2}}\sinh^{-1}\left(\frac{[L(L+1)g]^{\frac{1}{2}}}{[\eta^2+L(L+1)]^{\frac{1}{2}}}\right)\right\}$$

for $\rho<\eta$,

$$\left\{\begin{matrix}F_L(\rho)\\ G_L(\rho)/\sqrt{3}\end{matrix}\right\}=|\pi\varphi/6g^{\frac{1}{2}}|^{\frac{1}{2}}[-I_{-\frac{1}{3}}(|\varphi|)\pm I_{\frac{1}{3}}(|\varphi|)].$$

One may use the definition of Whittaker and Watson, namely that $I_{-\frac{1}{3}}-I_{\frac{1}{3}}=(2\sqrt{3}/\pi)K_{\frac{1}{3}}$, to write the result for F_L in an alternative form.

The Bessel functions of order one-third are not single valued in the vicinity of $x=0$, whereas the functions $x^{\frac{1}{2}}J_{\pm\frac{1}{3}}(2x^{\frac{1}{2}}/3)$ are indeed single-valued in this vicinity. In consequence it is convenient to utilize the latter functions, especially since they have been tabulated for complex values of their arguments.[19] Changing the notation to facilitate the use of these tables one has

$$\left\{\begin{matrix}F_L(\rho)\\ G_L(\rho)/\sqrt{3}\end{matrix}\right\}=\pm(3\pi/8)^{\frac{1}{2}}|\varphi|^{1/6}|g|^{-\frac{1}{4}}$$

$$\times\left[\left\{\begin{matrix}1\\ \frac{1}{3}\end{matrix}\right\}\mathrm{Re}\,h_1(z)\mp(\tfrac{1}{3})^{\frac{1}{2}}\,\mathrm{Im}\,h_1(z)\right], \quad (34)$$

where φ was defined earlier in Eqs. (33) and (33.1),

$$z=(\tfrac{3}{2})^{\frac{1}{2}}|\varphi|^{\frac{1}{2}}, \qquad \rho\geqslant\rho_L,$$

$$=-(\tfrac{3}{2})^{\frac{1}{2}}|\varphi|^{\frac{1}{2}}, \quad \rho\leqslant\rho_L.$$

For some applications, it is useful to have similar approximations for the derivatives $dF_L/d\rho$ and $dG_L/d\rho$ which are denoted, as customary, by F_L' and G_L'. While these derivatives follow immediately from the preceding formulae, the presence of absolute value signs (which were introduced as a convenience only) complicates the situation, and the explicit results are therefore given below,

$$\left\{\begin{matrix}F_L'(\rho)\\ G_L'(\rho)/\sqrt{3}\end{matrix}\right\}=[g^{\frac{1}{2}}/6|\varphi|-\tfrac{1}{2}d\ln(g^{\frac{1}{2}})/d\rho]\left\{\begin{matrix}F_L(\rho)\\ G_L(\rho)/\sqrt{3}\end{matrix}\right\}$$

$$+(\tfrac{2}{3})^{\frac{1}{2}}(3\pi/8)^{\frac{1}{2}}|\varphi|^{-1/6}g^{\frac{1}{2}}$$

$$\times\left[\left\{\begin{matrix}1\\ \frac{1}{3}\end{matrix}\right\}\mathrm{Re}\,h_1'(z)\pm(\tfrac{1}{3})^{\frac{1}{2}}\,\mathrm{Im}\,h_1'(z)\right], \quad (35)$$

where z is the same as in Eq. (34). It should be mentioned that $h_1'(z)$ is tabulated along with $h_1(z)$.

It is of interest to check the Wronskian of these approximations to F_L, G_L and F_L', G_L'. Employing the properties of $h_1(z)$ and $h_1'(z)$, one finds indeed that $F_L'G_L-G_L'F_L=1$ for the functions of Eqs. (34) and (35).

The Morse-Feshbach approximation to the Coulomb wave functions is not, in general, sufficiently accurate. To get some estimate of the accuracy of the fit one can examine the values assumed by the approximate solutions at $\rho=\rho_L$ and as $\rho\to\infty$, as well as $\rho\to 0$. At $\rho=\rho_L$ one finds

$$F_L(\rho_L)\to[\Gamma(\tfrac{1}{3})/2\pi^{\frac{1}{2}}]\left(\frac{\rho_L}{3[1+L(L+1)/\rho_L{}^2]}\right)^{1/6}.$$

This differs from the exact value [see Eq. (13)] in order $\rho_L{}^{-4/3}$ compared to unity. In order to improve the accuracy one may proceed in several ways:

(a) The difference between the exact potential and the potential in the differential equation satisfied by the approximate solution is

$$g(x)-z(dz/dx)^2=-\frac{(d^2/dx^2)[(dz/dx)^{-\frac{1}{2}}]}{(dz/dx)^{\frac{1}{2}}}\equiv\Delta(x).$$

Thus one can write

$$[d^2/dx^2+g(x)-\Delta(x)]\psi=-\Delta(x)\psi=\mathcal{O}\psi,$$

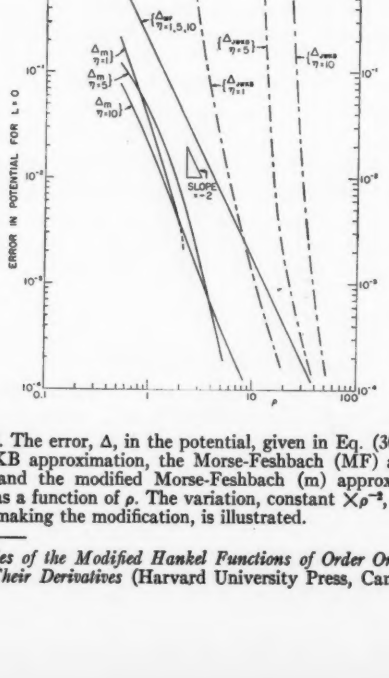


FIG. 8. The error, Δ, in the potential, given in Eq. (30.1), for the JWKB approximation, the Morse-Feshbach (MF) approximation and the modified Morse-Feshbach (m) approximation plotted as a function of ρ. The variation, constant $\times\rho^{-2}$, of Δ_{MF}, used in making the modification, is illustrated.

[19] *Tables of the Modified Hankel Functions of Order One-Third and of Their Derivatives* (Harvard University Press, Cambridge, 1945).

and invert this equation approximately to obtain

$$\psi \cong \psi_{M-F} + \mathcal{O}^{-1}[\Delta(x)\psi_{M-F}].$$

This procedure is closely allied to that used earlier to obtain Eqs. (24) and (25), but suffers from the complicated form of $\Delta(x)$, which all but precludes anything but numerical work.

(b) One may take a more complicated function for $\chi(z)$. In practice, this amounts to a quadratic fitting of the potential at every point (just as the Morse-Feshbach approximation was a linear fitting). While of value as a general technique, it is of little interest for the problem at hand since the resulting approximate solutions involve the parabolic cylinder functions, which are certainly less thoroughly tabulated than the Coulomb functions themselves.

(c) One may introduce in place of $g(x)$ an approximate function containing one or more arbitrary parameters and adjust the Morse-Feshbach approximation for this latter potential so as to minimize the error for the particular case at hand.

It turns out that procedure (c) is quite well suited to approximating the Coulomb functions, owing to the fortunate circumstance that the error in the potential, i.e., $\Delta(x)$, is, to a fair approximation, constant $\times\rho^{-2}$ in the region zero to infinity. Figure 8 illustrates this.

The application of this fact is immediate, since it involves no additional complication in the Morse-Feshbach approximation, requiring only a shift in the value of L. This is reminiscent somewhat of the shift $L\to L+\frac{1}{2}$ given by Kramers, and others,[5,18] for the JWKB approximation. The value of the shift is, however, not unique; one can choose it, for example, so as to make $\Delta(\rho)$ vanish for any specified value of ρ. [$\Delta(\rho)$ is, of course dependent on the altered value of L.] A reasonable place to make $\Delta(\rho)$ vanish is at the (new) turning point. For this, one finds that if

$$g^{\frac{1}{2}}=[1-2\eta/\rho-\alpha/\rho^2]^{\frac{1}{2}},$$

$$\Delta(\rho_a)=-\left(\frac{3}{140\eta^5}\right)\left(\frac{10+13\alpha/\eta^2-2(1+\alpha/\rho^2)^{\frac{1}{2}}}{(1+\alpha/\eta^2)\rho_a^2}\right),$$

where $\rho_a=\eta+(\eta^2+\alpha)^{\frac{1}{2}}$. Hence to cancel the error in the potential at ρ_a one must use

$$\alpha=L(L+1)+(3/140)\left(\frac{10+13\alpha/\eta^2-2(1+\alpha/\eta^2)^{\frac{1}{2}}}{(1+\alpha/\eta^2)}\right)$$

$$=L(L+1)+6/35+(3/35\eta^2)[L(L+1)+6/35]+\cdots.$$

The approximation $\alpha=L(L+1)+6/35$ is sufficient for η large.

To examine the usefulness of this artifice consider first $L=0$ in detail. The proposed approximation is, for $\rho\gtrsim\eta+(\eta^2+6/35)^{\frac{1}{2}}$, given by

$$g^{\frac{1}{2}}=[1-2\eta/\rho-6/35\rho^2]^{\frac{1}{2}},$$

$$\varphi=\rho g^{\frac{1}{2}}-\eta\ln[(\rho-\eta+\rho g^{\frac{1}{2}})/(\eta^2+6/35)^{\frac{1}{2}}]$$

$$-(6/35)^{\frac{1}{2}}\sin^{-1}[(6g/35)^{\frac{1}{2}}/(\eta^2+6/35)^{\frac{1}{2}}], \quad (36)$$

$$\left\{\begin{matrix}F_0(\rho)\\ G_0(\rho)/\sqrt{3}\end{matrix}\right\}=(\pi\varphi/6g^{\frac{1}{2}})^{\frac{1}{2}}[J_{-\frac{1}{3}}(\varphi)\pm J_{\frac{1}{3}}(\varphi)].$$

Consider the error near the new turning point, *i.e.*, at the point where $\rho=\eta+(\eta^2+6/35)^{\frac{1}{2}}\equiv\rho_a$. Using the results given by Eqs. (5) and (6) in a Taylor series, one readily finds that

$$F_0(\rho_a)_{\text{approx}}/F_0(\rho_a)_{\text{exact}}=1+1/3150\eta^2+\mathcal{O}(\eta^{-10/3}).$$

The error is surprisingly small, particularly in view of the fact that for the unmodified value of $L=0$, one found earlier that

$$F_0(2\eta)_{\text{approx}}/F_0(2\eta)_{\text{exact}}$$

$$=1+\eta^{-4/3}(\tfrac{3}{2})^{\frac{1}{3}}3\Gamma(\tfrac{2}{3})/35\Gamma(\tfrac{1}{3}).$$

Thus the error is pushed to order η^{-2} instead of order $\eta^{-4/3}$ and moreover the coefficient is remarkably small: $\approx 1/3000$ instead of $\approx 1/20$.

It is of interest to examine in more detail the origins of this accuracy, especially for general L. First of all the series for F_L and G_L had no terms of order $\eta^{-\frac{1}{3}}$, and F_L and G_L differed in the sign of the $\eta^{-4/3}$ term. Since the values for F_L and $G_L/\sqrt{3}$ given by the Morse-Feshbach approximation are the same at the turning point, the error in this approximation is therefore of order $\eta^{-4/3}$, already one higher order than might be expected. Now the use of the modified value of L shifts the value of the turning point to

$$\rho_a=\rho_0+\eta^{-1}[3/35+L(L+1)/2]$$

$$+\eta^{-3}[9/2450-L^2(L+1)^2/8]+\mathcal{O}(\eta^{-5}).$$

[This uses the modified value for $L(L+1)$ to be $L(L+1)+6/35+(3/35\eta^2)(L(L+1)+6/35)+\cdots$.] To obtain the value of $F_L(\rho_a)$ and $G_L(\rho_a)$ one uses a Taylor series

$$F_L(\rho_a)=F_L(\rho_0)+(1/2\eta)(6/35+L(L+1))F_L'(\rho_0)+\cdots.$$

Upon referring to Eqs. (5) and (6), it is clear that the shift from ρ_L to ρ_a is precisely that required to cancel terms of order $\eta^{-4/3}$, for both F_L and G_L. In fact, the requirement that the shift minimize the error in the wave functions seems equally as good a criterion as the equivalent requirement that the error in the potential cancel at ρ_a. The terms of order η^{-2} do not cancel and,

for general L, one has

$$\begin{Bmatrix} F_L(\rho_a) \\ G_L(\rho_a)/\sqrt{3} \end{Bmatrix} = \left(\frac{\Gamma(\frac{1}{3})}{2\pi^{\frac{1}{2}}}\right)(\tfrac{2}{3}\eta)^{1/6} \times\left[1-\left(\frac{1}{3150}\right)\eta^{-2}\mp\left(\frac{216a}{398\,125}\right)\eta^{-10/3}+\cdots\right], \quad (37)$$

where $a=(\frac{3}{2})^{\frac{1}{3}}\Gamma(\frac{2}{3})/\Gamma(\frac{1}{3})$.

This result requires η to be large and L/η to be small. Note the remarkable independence of L shown in this result.

Now the value assumed by the modified Morse-Feshbach approximation at ρ_a is the same for both F_L and G_L and turns out to be

$$\begin{Bmatrix} F_L(\rho_a) \\ G_L(\rho_a)/\sqrt{3} \end{Bmatrix} \cong \left(\frac{\Gamma(\frac{1}{3})}{2\pi^{\frac{1}{2}}}\right)(\tfrac{2}{3}\eta)^{\frac{1}{2}}[1+\mathcal{O}(\eta^{-4})]. \quad (37.1)$$

Hence one obtains the result that, for general L, the modified Morse-Feshbach approximation yields

$$\begin{Bmatrix} F_L(\rho_a)_{\text{approx}}/F_L(\rho_a)_{\text{exact}} \\ G_L(\rho_a)_{\text{approx}}/G_L(\rho_a)_{\text{exact}} \end{Bmatrix} = 1+\left(\frac{1}{3150}\right)\eta^{-2} \pm\left(\frac{216a}{398\,125}\right)\eta^{-10/3}+\mathcal{O}(\eta^{-4}), \quad (38)$$

which, to repeat, is both gratifyingly accurate and moreover independent of L for L/η small.

For $\rho\to\infty$ the results are also very good. As is typical of the Morse-Feshbach approximation, the approximate result differs asymptotically from the exact result only in the asymptotic phase. That is,

$$F_L(\rho)\sim\sin(\rho-L\pi/2-\eta\ln 2\rho+\sigma_L),$$

and the approximate result differs from this exact result only by a different value for σ_L, namely, for $L=0$,

$$(\sigma_0)_{\text{approx}}=\pi/4-\eta+\eta\ln\eta+(\eta/2)\ln(1+6/35\eta^2) -(6/35)^{\frac{1}{2}}\sin^{-1}[1/(1+35\eta^2/6)^{\frac{1}{2}}]. \quad (39)$$

Expanding in inverse powers of η with the help of Sterling's series for the exact phase yields the result

$$(\sigma_0)_{\text{approx}}=(\sigma_0)_{\text{exact}}-1/420\eta+\mathcal{O}(\eta^{-3}).$$

The unmodified Morse-Feshbach result is

$$(\sigma_0)_{\text{approx}}=(\sigma_0)_{\text{exact}}+1/12\eta+\mathcal{O}(\eta^{-3}).$$

Hence the modification results in a significant improvement.

For general L, the value for σ_L which results from the modification is

$$(\sigma_L)_{\text{approx}}=(\eta/2)\ln[\eta^2+L(L+1)+6/35]-\eta +(L+\tfrac{1}{2})(\pi/2)-[L(L+1)+6/35]^{\frac{1}{2}} \times\sin^{-1}\left[1\Big/\left(1+\frac{\eta^2}{[L(L+1)+6/35]}\right)^{\frac{1}{2}}\right], \quad (39.1)$$

[where $L(L+1)$ has been replaced by $L(L+1)+6/35$].

For large η, with L/η small, this becomes

$$(\sigma_L)_{\text{approx}}=(\sigma_L)_{\text{exact}}-1/420\eta+\mathcal{O}(\eta^{-3}),$$

which is, once again, independent of L to the order given.

Finally these approximate wave functions can be examined for $\rho\to 0$. Using the asymptotic form for $h_1(z)$ for $z\to-\infty$ one finds

$$\begin{Bmatrix} 2F_L(\rho) \\ G_L(\rho) \end{Bmatrix}_{\text{approx}} \to|g|^{-\frac{1}{2}}e^{\mp|\varphi|}[1\mp 5/72|\varphi|+\cdots], \quad (40)$$

$$|\varphi|\to[L(L+1)+6/35]^{\frac{1}{2}}[\ln\rho+\mathcal{O}(\rho^0)],$$

as $\rho\to 0$. Thus

$$\begin{Bmatrix} 2F_L(\rho) \\ G_L(\rho) \end{Bmatrix}\to\rho^{f(L)}, \quad (40.1)$$

where $f(L)=(\frac{1}{2})\pm[L(L+1)+6/35]^{\frac{1}{2}}\ln\rho+\mathcal{O}(\rho^0)$.

The exponents for ρ, for the exact F_L and G_L, should be, of course, $L+1$ and $-L$. The approximation yields however, $f(L)\sim\frac{1}{2}+[L(L+1)+6/35]^{\frac{1}{2}}$ with the upper sign, and $f(L)\sim\frac{1}{2}-[L(L+1)+6/35]^{\frac{1}{2}}$ with the lower.

For moderately large L the square root can be expanded and one finds

$$f(L)\sim\tfrac{1}{2}\pm(L+\tfrac{1}{2}+\tfrac{1}{2}(6/35-\tfrac{1}{4})[1/(L+\tfrac{1}{2})+\cdots] =\tfrac{1}{2}\pm(L+\tfrac{1}{2})\mp 11/140(2L+1)+\cdots. \quad (39.1)$$

The error in the exponents decreases reasonably well as L increases.

For $L=0$ one has as exponents $\frac{1}{2}+(6/35)^{\frac{1}{2}}\cong 32/35$ and $\frac{1}{2}-(6/35)^{\frac{1}{2}}\cong 3/35$, which are to be compared to 1 and 0 respectively. The approximation to F_0 is satisfactory, but for G_0 the approximation is very poor for extremely small ρ. Nevertheless, the 6/35 modification did result in marked improvement.

The consideration of the special values of ρ above, namely $\rho=0$, $\rho_a=\infty$, indicate clearly that the proposed modification of the Morse-Feshbach approximation yields a very satisfactory approximation to the Coulomb wave functions over the entire positive real axis. This is, of course, largely due to the simplicity of g for the Coulomb case, since only one turning point occurs for $\rho\geqslant 0$. The results obtained for the approximate wave functions using these special values of ρ are summarized in Table I for convenience.

TABLE I. Approximate wave functions for special values of ρ.

Exponents as $\rho\to 0$	$\begin{Bmatrix}(F_L)_{\text{approx}}/(F_L)_{\text{exact}}\\(G_L)_{\text{approx}}/(G_L)_{\text{exact}}\end{Bmatrix}$ at turning point	Error in σ_L at $\rho=\infty$
MF: $\frac{1}{2}$, $\frac{1}{2}$	$1\pm(3/35)(\frac{3}{2})^{1/3}\eta^{-4/3}\Gamma(\frac{2}{3})/\Gamma(\frac{1}{3})+\cdots$	$1/12\eta$
MF mod: $\frac{1}{2}+[6/35+L(L+1)]^{\frac{1}{2}}$, $\frac{1}{2}-[6/35+L(L+1)]^{\frac{1}{2}}$	$1+1/3150\eta^2+\cdots$	$-1/420\eta$
Exact: $L+1$, $-L$	1	0

In order to illustrate the accuracy of these approximations for other than the special values of ρ used above sample calculations have been performed for F_0 and G_0 which are tabulated in Table II. It should be noted that even for $\eta=1$, which is a severe test of the approximations, the agreement is good. The error decreases rapidly as η increases. The few comparisons made reflect primarily the lack of tables for large η.

In the discussion above for F_L and G_L at the turning point, the expansions have been made under explicit assumption that $\eta\gg 1$ and L/η is small. It is of some interest to note that the approximate wave functions are equally valid for L/η not small. At the modified turning point, ρ_a, it has already been found that

$$\begin{Bmatrix}F_L(\rho_a)\\G_L(\rho_a)/\sqrt{3}\end{Bmatrix}\cong[\Gamma(\tfrac{1}{3})/2\pi^{\frac{1}{2}}]6^{-1/6}\rho_a^{\frac{1}{2}}(\rho_a-\eta)^{-1/6}.$$

The definition of ρ_a is, in general,

$$\rho_a=\eta+(\eta^2+\alpha)^{\frac{1}{2}},$$

where α is the modified value of $L(L+1)$ as given earlier. For $L/\eta\gg 1$ one has therefore

$$\alpha=L(L+1)+39/140+\mathcal{O}(1/L).$$

As a result, one finds that

$$\rho_a=\rho_L+39/280\rho_L+\mathcal{O}(\rho_L^{-2}).$$

Hence the Morse-Feshbach (modified) approximation to F_L and G_L at ρ_a assumes the form

$$\begin{Bmatrix}F_L\\G_L/\sqrt{3}\end{Bmatrix}\cong[\Gamma(\tfrac{1}{3})/2\pi^{\frac{1}{2}}]6^{-1/6}\rho_L^{1/6}\times[1+\eta/6\rho_L+\mathcal{O}(1/\rho_L^2)]. \quad (41)$$

To obtain the exact value for F_L and G_L at ρ_a, one again resorts to a Taylor series. Noting that $F_L''(\rho_L)$ and $G_L''(\rho_L)$ vanish, it is found that

$$\begin{Bmatrix}F_L(\rho_a)\\G_L(\rho_a)\end{Bmatrix}=\begin{Bmatrix}F_L(\rho_L)\\G_L(\rho_L)\end{Bmatrix}+(39/280\rho_L)\begin{Bmatrix}F_L'(\rho_L)\\G_L'(\rho_L)\end{Bmatrix}+\mathcal{O}(\rho_L^{-4})\begin{Bmatrix}F_L\\G_L\end{Bmatrix}.$$

Under the assumptions that $L\gg\eta\gg 1$, Eqs. (13) and (14) assume the forms,

$$\begin{Bmatrix}F_L(\rho_L)\\G_L(\rho_L)/\sqrt{3}\end{Bmatrix}\to(\Gamma(\tfrac{1}{3})/2\pi^{\frac{1}{2}})\rho_L^{1/6}6^{-1/6}(1-\eta/\rho_L)^{-1/6}\times(1\mp 39b2^{\frac{1}{2}}\rho_L^{-4/3}/280+\cdots),$$

$$\begin{Bmatrix}F_L'(\rho_L)\\G_L'(\rho_L)/\sqrt{3}\end{Bmatrix}\to[\Gamma(\tfrac{1}{3})/2\pi^{\frac{1}{2}}]\rho_L^{1/6}6^{-1/6}\times[\pm b\rho^{-\frac{1}{2}}2^{\frac{1}{2}}+(3(2)^{-4/3}/5b)\rho_L^{-1}+\cdots].$$

Thus,

$$\begin{Bmatrix}F_L(\rho_a)\\G_L(\rho_a)/\sqrt{3}\end{Bmatrix}=[\Gamma(\tfrac{1}{3})/2\pi^{\frac{1}{2}}]\rho_L^{1/6}6^{-1/6}(1-\eta/\rho_L)^{-1/6}\times[1+\mathcal{O}(\rho_L^{-2})]. \quad (41.1)$$

This result is seen to agree with the approximate result, Eq. (41), up to order ρ_L^{-2}. It is clear therefore that the approximate wave functions give excellent results for F_L and G_L at $\rho=\rho_a$ in both limits $L\gg\eta$ and $\eta\gg L$. This leads one to feel confident that similarly good results will hold for the transition region, $L\sim\eta$, as well. It is important to note, however, that the aforementioned results all require the use of a value for the "shifted" $L(L+1)$, that is α, which is appropriate to the region of interest. Now α can be given quite generally, but this would require solving a quartic equation, given earlier. The result is quite unwieldy, and it was therefore considered reasonable to confine attention to limiting cases as done above.

TABLE II. Sample calculations for F_0 and G_0.

$\eta;\rho$	$F_0(\rho)$ JWKB	MF	MF mod.	Exact	$G_0(\rho)$ JWKB	MF	MF mod.	Exact
1; 0.6	0.1393	0.1322	0.1177	0.1071	1.644	2.564	2.631	2.792
1; 3	1.2268	1.1111	1.0877	1.0844	0.4763	0.5591	0.6220	0.6284
1; 6	−0.2742	−0.2465	−0.1603	−0.1665	−1.072	−1.074	−1.096	−1.090
1.995; 1.2	0.04853	0.04725	0.04370	0.04343	6.756	7.0747	7.4526	7.5056
3.981; 2.4	0.005923	0.005834	0.005635	0.005625	5.5450	56.544	58.162	58.280

APPENDIX

For L of the same order as or larger than η, F_L and G_L are not as similar to each other at ρ_0 as they are at ρ_L, and if L is quite large compared to η, the difference may be very great. In this case it is necessary to turn to the individual integral representations of F_L and G_L as given in reference 13, Eqs. (10.3) and (10.5)

$$F_L(\rho)=\frac{e^{-\pi\eta}\rho^{L+1}}{(2L+1)!C_L}\int_0^1(1-z^2)^L\cos(2\eta\tanh^{-1}z-\rho z)dz, \tag{A1}$$

$$G_L(\rho)=\frac{e^{-\pi\eta}\rho^{L+1}}{(2L+1)!C_L}\left\{\int_0^1(1-z^2)^L\sin(2\eta\tanh^{-1}z-\rho z)dz\right.$$
$$\left.+e^{\pi\eta}\int_0^\infty(1+u^2)^L\exp[-u\rho-2\eta\tan^{-1}(1/u)]du\right\}. \tag{A2}$$

To obtain $F_L(\rho_0)$, note that the main contribution to the integral comes from small z and expand the cosine for z small, setting $\rho=2\eta$ wherever it occurs. Term-by-term integration, together with the use of the expansion of C_L already discussed, yields the following result for $L\gtrsim\eta$:

$$F_L(\rho_0)\cong[(1/2L)^{\frac{1}{2}}\eta^{L+1}(L^2+\eta^2)^{-(L/2+\frac{1}{4})}/(1+1/2L)]$$
$$\times\exp\{L-\eta\tan^{-1}L/\eta-5\eta^2/12L^3-1/8L$$
$$-L/12(L^2+\eta^2)\}(1+11\eta^2/8L^4+5\eta^4/4L^6+\cdots). \tag{A3}$$

The expansion for G_L at ρ_0 may be obtained in terms of the steepest descents result, given in reference 13, Eq. (9.6), for $\rho=2\eta$. For large η, only the second integral in Eq. (A2) contributes, and corrections to their Eq. (9.6) may be obtained by expanding the integrand about $u=L/\eta$, and integrating term-by-term. The result, after treating the coefficient in the same manner as before, is

$$G_L(\rho_0)=(2/L)^{\frac{1}{2}}\eta^{-L}(L^2+\eta^2)^{(L/2+\frac{1}{4})}$$
$$\times\exp\{-L+\eta\tan^{-1}L/\eta+5\eta^2/12L^3-1/8L$$
$$+L/12(L^2+\eta^2)\}(1+\eta^2/8L^4+5\eta^4/4L^6). \tag{A4}$$

With the help of the recurrence relation given by Powell,[20] the derivatives of $F_L(\rho_0)$ and $G_L(\rho_0)$ were found to be

$$F_L'(\rho_0)=F_L(\rho_0)\{[(L+1)/2\eta][1+\eta^2/L^3$$
$$-7\eta^2/2L^4-5\eta^4/2L^6]\},$$
$$G_L'(\rho_0)=G_L(\rho_0)\{-[L/2\eta][1-\eta^2/L^3 \tag{A5}$$
$$-\eta^2/2L^4-5\eta^4/2L^6]\}.$$

In order to obtain an expansion of $F_L(\rho)$ for arbitrary ρ in the vicinity of ρ_0, one must return to the integral representation of Eq. (A1) and, as before, expand for z small. In the present case, an extra term $(\rho_0-\rho)z$ appears as part of the argument of the cosine in the integrand, and necessitates use of the expression for the cosine of the sum of two angles. Noting that

$$\int_0^1(1-z^2)^Lz^{2m}\cos(\Delta\rho z)dz=L!\pi^{\frac{1}{2}}2^{L-\frac{1}{2}}(-1)^m$$
$$\times D^{2m}[(\Delta\rho)^{-(L+\frac{1}{2})}J_{L+\frac{1}{2}}(\Delta\rho)]\equiv b_{2m},$$

where $D\equiv d/d\rho$, and

$$\int_0^1(1-z^2)^Lz^{2m+1}\sin(\Delta\rho z)dz=L!\pi^{\frac{1}{2}}2^{L-\frac{1}{2}}(-1)^m$$
$$\times D^{2m+1}[(\Delta\rho)^{-(L+\frac{1}{2})}J_{L+\frac{1}{2}}(\Delta\rho)]\equiv b_{2m+1},$$

where $\Delta\rho=\rho_0-\rho$, one may write the resulting expression for $F_L(\rho)$, when $L\gtrsim\eta$ and $(\rho_0-\rho)$ small, as

$$F_L(\rho)=[e^{-\pi\eta}\rho^{L+1}/(2L+1)!C_L]$$
$$\times\{b_0-2\eta(b_3/3+b_5/5+b_7/7+\cdots)$$
$$-2\eta^2(b_6/9+2b_8/15+\cdots)+\eta^3b_9/81+\cdots\}. \tag{A6}$$

The corresponding expression for $G_L(\rho)$ is again obtained by finding corrections to the steepest descents approximation. The result is

$$G_L(\rho)=[G_L(\rho)]_{SD}\{1+1/24L+[(L^2+\eta^2)\rho^2/24L]$$
$$\times[5L/(X-L)^3(LX+\rho\eta)$$
$$-X^2/(X-L)^2(LX+\rho\eta)^2]+\cdots, \tag{A7}$$

where, according to Eq. (9.6) of reference 13,

$$[G_L(\rho)]_{SD}=2^L(2\pi)^{\frac{1}{2}}(LX+\rho\eta)^{L+\frac{1}{2}}$$
$$\times\exp\{-X-2\eta\tan^{-1}(\rho/X)\}/$$
$$(2L+1)!C_L\rho^L(X-L)^{\frac{1}{2}}, \tag{A8}$$

and $X=L+(L^2+2\rho\eta-\rho^2)^{\frac{1}{2}}$.

[20] J. L. Powell, Phys. Rev. 72, 626 (1947).

Letters to the Editor

PUBLICATION of brief reports of important discoveries in physics may be secured by addressing them to this department. The closing date for this department is five weeks prior to the date of issue. No proof will be sent to the authors. The Board of Editors does not hold itself responsible for the opinions expressed by the correspondents. Communications should not exceed 600 words in length and should be submitted in duplicate.

Quenching-In of Lattice Vacancies in Pure Gold*

J. W. KAUFFMAN AND J. S. KOEHLER
University of Illinois, Urbana, Illinois
(Received November 22, 1954)

PRELIMINARY data obtained by the quenching of 99.999 percent pure gold wires was reported in an earlier letter.[1] The quantities of interest are E_F, the energy required to produce a lattice vacancy; and E_M, the activation energy required for vacancy motion. The present letter reports on more accurate measurements.

E_F is measured by determining the increases in resistance produced by quenching a specimen from various temperatures in the range from 690°C to 900°C. One has $\Delta R = Ae^{-E_F/kT}$, where A is a constant and ΔR is the increase in the resistance at liquid nitrogen temperature produced by a quench from the high temperature T. Both 16-mil and 25-mil wires were used and the quench by a precooled jet of helium gas from the high temperature to room temperature occurred in about 10 milliseconds. The wire was then manually turned down into liquid nitrogen (time required, about a third of a second). In the case of the 25-mil wires we were unable to secure a sufficiently rapid cooling to retain all of the vacancies on quenching from temperatures above 800°C. The data from wires of both sizes showed the exponential increase in quenched in resistance but the data obtained with the 16-mil wires gave the more accurate values of E_F. The limiting low temperature which can be used for quenching appears to be determined by the amount of impurity present. The residual resistance of many of the specimens was measured at liquid helium temperature. The values obtained indicated that the wires were about 99.999 percent pure. It was noted that on specimens of rather low purity two difficulties arise: First, faster quenching rates are required to retain all of the imperfections. Second, on quenching from various temperatures below some temperature T_0 one finds that very nearly the same resistance ΔR_0 is introduced. Both T_0 and ΔR_0 increase with increasing amounts of impurity thus limiting the temperature range over which a determination of E_F can be made. No cold-work was introduced because all of the resistance increase could be annealed out by a 600°C anneal. Permanent changes in dimensions would be introduced by cold-work.

Since liquid nitrogen dissolves oxygen which changes the boiling temperature a similar dummy specimen was also measured to correct for changes in the bath temperature. R was measured to better than one part in thirty thousand.

It was found that $E_F = 1.28 \pm 0.03$ ev for gold. A quench from 816°C produced an increase of 0.21×10^{-8} ohm cm in the resistivity.

Annealing measurements were made to determine the activation energy of motion E_M for lattice vacancies in gold. Isothermal annealing measurements were made in which R was measured at liquid nitrogen temperature after various lengths of time in a constant temperature bath. At a certain stage in the annealing, the temperature of the annealing bath would be increased. The increase in the rate of annealing with increase in annealing temperature at constant R depends on E_M as follows:

$$\frac{dR_2}{dt} \Big/ \frac{dR_1}{dt} = \exp\left[-E_M\left(\frac{1}{kT_2} - \frac{1}{kT_1}\right)\right].$$

The annealing baths were operated between -30°C and $+15$°C. Their temperature was constant to ± 0.01°C. It was found that $E_M = 0.68 \pm 0.03$ ev. About half of the quenched-in resistivity recovered after 57 hours at 25°C.

If self diffusion in gold occurs by means of vacancies, then it can be shown that the activation energy for self diffusion, Q, should be given by: $Q = E_M + E_F$. The above data give $Q = 1.96 \pm 0.06$ ev. Gatos and Kurtz[2] have recently measured Q for gold, and find $Q = 1.965$ ev. The agreement is most satisfactory.

* This work is supported by the U. S. Atomic Energy Commission.

[1] J. W. Kauffman and J. S. Koehler, Phys. Rev. 88, 149 (1952).
[2] H. C. Gatos and A. D. Kurtz, J. Metals 6, 616 (1954).

Ferromagnetic Resonance in Nickel Ferrite Between One and Two Kilomegacycles

H. SUHL
Bell Telephone Laboratories, Inc., Murray Hill, New Jersey
(Received November 24, 1954)

DUE to limitations imposed by internal anisotropy fields, microwave resonance work in ferrites has largely been restricted to frequencies above 4 kMc/sec. This note describes special experimental arrangements, and observations on one of these, which make it possible to compound an applied field with the internal shape and crystalline anisotropy field to yield a particularly low effective field, and thus a low resonant frequency.

The method involved is best demonstrated for a (cubic) single-crystal, single-domain ferrite sphere, with cube axes and diagonals the hard and easy magnetization directions respectively. In zero applied field the

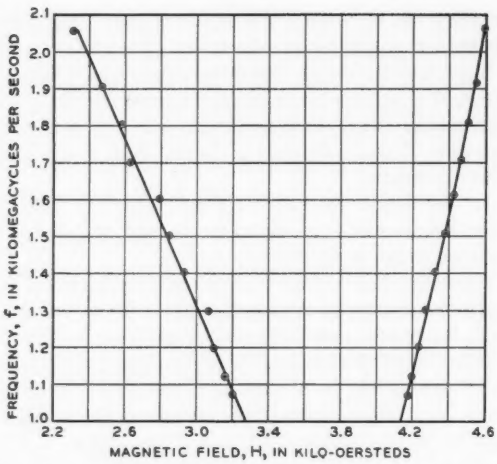


FIG. 1. Experimental curve of resonant frequency *vs* field, showing the characteristic two-arm structure.

magnetization M is aligned along 111, say, a minimum in the energy surface. Application of a dc field H in the (hard) 001 direction pulls the energy minimum towards 001 making it more shallow in the process. The natural precession frequency of M about the minimum is thus decreased. Finally at a certain critical field H_{crit}, the minimum reaches the maximum at 001, yielding a local "flat-spot" in the energy surface. M now has no "stiffness" at all, the natural precession frequency is zero. Further increase in H leaves M aligned along 001, and simply gives a resonance frequency $\gamma(H-H_{crit})$, with γ the gyromagnetic ratio. H_{crit} for the sphere is $H_a=-2k_1/M$, where k_1 is the first anisotropy constant (here negative).

For the $Ni_{0.75}Fe_{2.25}O_4$ crystal investigated,[1] ($H_a\sim240$ oe, $M\sim290$ gauss) spherical geometry is unsuitable. Single domain behavior results only if the effective field for domain wall motion [about $H-(4/3)\pi M$ in the sphere] exceeds zero, a condition which fails near $H=H_{crit}$. However for a disk cut in the 001 plane, the same principles apply: at a certain critical field (now $4\pi M+H_a$) along 001 the resonant frequency goes to zero. The single domain condition (now $H-4\pi M>0$) is readily satisfied near H_{crit}. Measurements on a disk 0.2 in. in diameter and 0.005 in. thick are shown in Fig. 1.

More general arrangements are best examined by writing the total energy E (in external field, demagnetizing field, and anisotropy field) and the Landau-Lifschitz equations of motion in the polar coordinates θ, φ of the M vector. A small precession about equilibrium ($\partial E/\partial\theta=\partial E/\partial\varphi=0$) is then seen to have a natural frequency

$$\omega_{res}=\frac{\gamma}{M\sin\theta}\left[\frac{\partial^2E}{\partial\theta^2}\frac{\partial^2E}{\partial\varphi^2}-\left(\frac{\partial^2E}{\partial\theta\partial\varphi}\right)^2\right]^{\frac{1}{2}}.$$

This expression goes to zero whenever the energy surface becomes "parabolic" near the minimum,[2] a condition attainable in many ways. In the disk just described, it is attained for example with H anywhere in the 010 or 100 planes; H_{crit}, and its inclination θ_H to 001 being given parametrically in terms of θ:

$$H_{crit}^2=H_a^2\sin^6\vartheta+(H_d+H_a\cos^2\theta)^2\cos^2\theta,$$

$$\tan\theta_H=\frac{H_a\sin^2\theta}{H_d+H_a\cos^2\theta}\tan\theta,\quad H_d=4\pi M.$$

With H in the $1\bar{1}0$ plane, zero frequency is not always achieved. The curves computed in Fig. 2 have finite minima when H is inclined slightly into the octant occupied by M. Such misalignment is the more serious the greater M: with $H_d\sim3000$ oe, it need only be 30′ to give $\omega_{min}\sim700$ Mc/sec. Similar considerations presumably apply to other low symmetry orientations of H. Thus crystal lineage may be responsible for the apparent failure of the two arms of the curve in Fig. 1 to extrapolate to zero frequency. Also, some residual domain structure (especially near the rim) may prevent uniform lineup of M.

Detailed theory predicts the following features of the line profile: On the high-field branch of the curve in Fig. 1, the line-strength should be independent of rf magnetic field orientation in the plane of the disk; but on

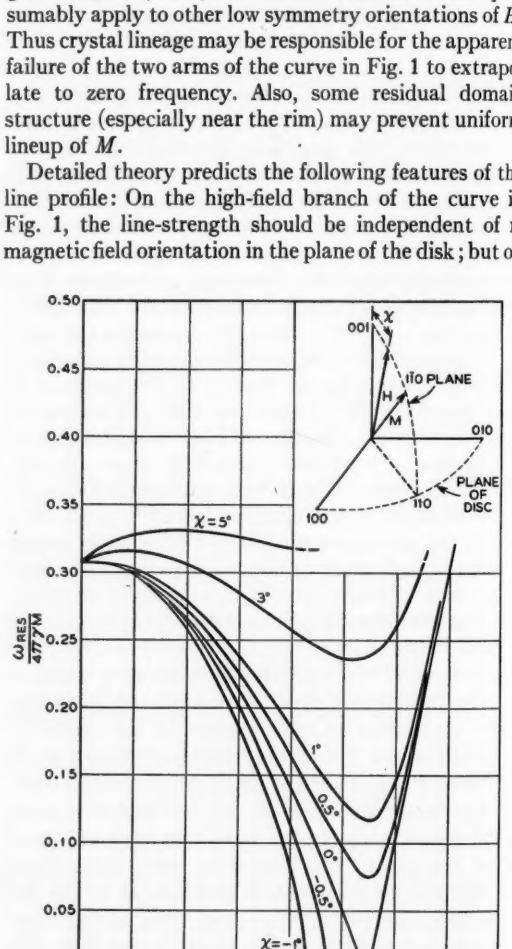


FIG. 2. Computed effect of misalignment of H in the 110 plane, for $H_a/H_d=0.1$. For $\chi<0$, the upper branch is unstable. k_2 has been neglected.

the lower branch the absorption should equal a constant term, plus a term varying as the sine squared of the angle between 110 and the rf field; thus as the disk is rotated in its own plane, there should be two absorption maxima 180° apart. The conclusions were well substantiated by experiment, except at the low-frequency end, where the upper branch also showed some sensitivity. The agreement indicates one, or at most two domains (in directions of equal latitude, with longitudes 180° apart), on the low field-branch. Four domains 90° apart in longitude would have given four absorption maxima.

The formula for linewidth (between points of $\frac{1}{2}$ maximum absorption on either side of resonance) is given by

$$\Delta H=\frac{\alpha}{|d\omega_{\rm res}/dH|}\frac{\gamma}{M}\left(E_{\theta\theta}+\frac{1}{\sin^2\theta}E_{\varphi\varphi}\right),$$

in terms of the Landau-Lifschitz loss parameter α. The α measured at 1920 Mc/sec on the upper branch of the curve was about twice that found in similar material at 24 kMc/sec.[3] This increase is yet unaccounted for.

The author is indebted to Dr. A. M. Clogston for suggesting the experiment, to Dr. L. R. Walker for his continued interest in the theory, to Dr. J. F. Dillon, Jr. for supervising the sample preparation, to Mr. J. H. Rowen for valuable advice on instrumentation, to Mrs. A. Rebarber for the computations, and to Mr. J. Davis for assistance in the experiment. The single crystal was supplied by Dr. W. Clarke of Linde Air Products.

[1] Composition approximate.
[2] This formula was also derived, independently, by J. Smit, (Talk at Conference on Ferrimagnetism at the Naval Ordnance Laboratory, 11–12 Oct., 1954), and by P. Tannenwald and B. Lax (private communication). The latter have done theoretical work of the kind reported here.
[3] Galt, Yager, and Merritt, Phys. Rev. **93**, 1119 (1954).

Ultrasonic Attenuation Due to Lattice-Electron Interaction in Normal Conducting Metals

W. P. Mason
Bell Telephone Laboratories, Murray Hill, New Jersey
(Received November 12, 1954)

In a recent letter to the editor,[1] Bömmel published some experimental results on the attenuation of sound waves at ultrasonic frequencies for single lead crystals, which showed that there was an increase in attenuation at very low temperatures for the normal conducting state which disappeared in the superconducting state. This attenuation difference occurred for both shear and longitudinal waves and increased in proportion to the square of the frequency. From 1.6°K to 4°K the difference was independent of the temperature and was 0.106 neper per cm for longitudinal waves of 26.65 Mc/sec and 0.061 neper per cm for shear waves of 9.5 Mc/sec. Figure 1 shows complete measurements

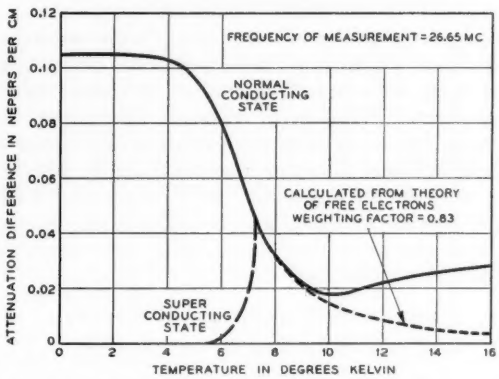

FREQUENCY OF MEASUREMENT = 26.65 MC
NORMAL CONDUCTING STATE
CALCULATED FROM THEORY OF FREE ELECTRONS WEIGHTING FACTOR = 0.83
SUPER CONDUCTING STATE
ATTENUATION DIFFERENCE IN NEPERS PER CM
TEMPERATURE IN DEGREES KELVIN

FIG. 1. Comparison of measured attenuation of a lead single crystal with that calculated from free electron theory.

for longitudinal waves. Bömmel has recently measured the same effect for a single tin crystal with the results shown by Fig. 2 for a longitudinal wave of 28.5 Mc/sec.

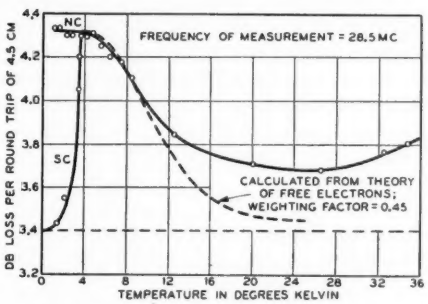

NC
FREQUENCY OF MEASUREMENT = 28.5 MC
SC
CALCULATED FROM THEORY OF FREE ELECTRONS; WEIGHTING FACTOR = 0.45
DB LOSS PER ROUND TRIP OF 4.5 CM
TEMPERATURE IN DEGREES KELVIN

FIG. 2. Comparison of measured attenuation in tin with that calculated from free electron theory.

A curve of similar shape is obtained for shear waves with a value of 0.036 neper per cm at 17 Mc/sec and 1.5°K for the difference between normal and superconducting states.

It is the purpose of this note to point out that a simple phenomenological concept of the interaction between the lattice vibrations and the electron gas gives values of attenuations which agree well with the measured results. The concept considered is that in the normal state a lattice vibration can communicate energy to the electron gas by a viscous reaction, i.e., transfer of momentum, and is damped by the viscosity of the gas, while in the superconducting state the lattice is not able to transfer momentum to the electron gas and the damping disappears.

For the most general case the attenuations[2] caused by the energy loss due to the shear and compressional viscosities of the electron gas are for longitudinal and

shear waves in the lattice

$$A_l(\text{nepers/cm})=\frac{2\pi^2 f^2}{\rho v_l^3}\left(\frac{4}{3}\eta+\chi\right);\quad A_s=\frac{2\pi^2 f^2}{\rho v_s^3}\eta, \quad (1)$$

where ρ is the density of the crystal, f is the frequency, η the shear viscosity, χ the compressional viscosity, v_l and v_s the velocities of longitudinal and shear waves respectively. This concept accounts directly for the increase in attenuation with the square of the frequency. The velocities of the waves were measured by Bömmel and were respectively $v_l=2.35\times10^5$ cm/sec, $v_s=1.266\times10^5$ cm^2/sec for lead and $v_l=3.48\times10^5$ cm/sec, $v_s=1.9\times10^5$ cm/sec for tin at 1.5°K. With these values the viscosities calculated are

$$\begin{aligned}\text{Lead: } &\eta=0.787 \text{ poise}; \quad \chi=0.05 \text{ poise};\\ \text{Tin: } &\eta=0.314 \text{ poise}; \quad \chi=0.03 \text{ poise}.\end{aligned} \quad (2)$$

The values of compressional viscosity are small and are probably within the experimental error, nearly zero.

To see if these values are reasonable for the viscosity of an electron gas, we make use of the formula for viscosity[3]:

$$\eta=Nml\bar{v}/3, \quad (3)$$

where N is the number of electrons per cc, m their mass, l the mean free path (in this case between electrons and lattice atoms) and $\bar{v}$ the mean velocity. These last two quantities can be evaluated from the theory of the free electron gas[4] and are

$$\bar{v}^2=\frac{3}{5}\frac{\hbar^2}{m^2}(3\pi^2N)^{\frac{1}{3}};\quad l=\frac{\sigma m\bar{v}}{Ne^2}, \quad (4)$$

where e is the charge on the electron, $\hbar$ is Planck's constant divided by 2π, and σ is the electrical conductivity in cgs units. Introducing (4) in (3) and substituting $\sigma=9\times10^{11}/R$, where R is the resistivity in ohm cm, the value of η becomes

$$\eta=9\times10^{11}\hbar^2(3\pi^2N)^{\frac{1}{3}}/(5e^2R). \quad (5)$$

The value of N for monovalent metals is equal to the number of atoms, which for lead and tin are respectively 3.33×10^{22} and 3.72×10^{22} per cc. For quadrivalent metals no adequate theory exists for correlating the number of electrons with the number of atoms, but experiments[5] on tin indicate that the ratio is between 0.43 and 1.1. All the other values are known and the viscosity is determined in terms of the measured resistivity by the equations

$$\begin{aligned}\eta_l&=(N_e/N_a)^{\frac{1}{3}}\times8.6\times10^{-9}/R;\\ \eta_t&=(N_e/N_a)^{\frac{1}{3}}\times9.2\times10^{-9}/R.\end{aligned} \quad (6)$$

The measured resistivities[6] for lead and tin are shown in Fig. 3. Using the values of Fig. 3 and Eq. (6) to determine the viscosity, and substituting in Eq. (1) for a longitudinal wave, the calculated attenuations are shown by the dashed lines. Best agreements are obtained if the factors N_e/N_a are

$$N_e/N_a=0.75, \text{ lead}; \quad N_e/N_a=0.3, \text{ tin}.$$

These values account well for the shapes of the measured curves. Above 10°K additional loss occurs due to dislocation motions.

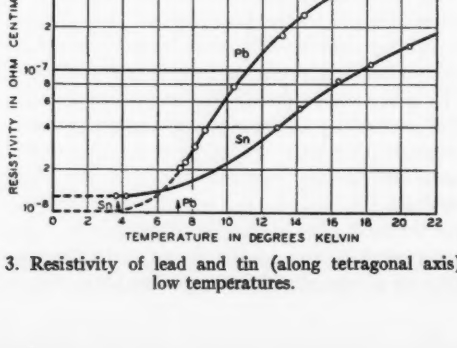


FIG. 3. Resistivity of lead and tin (along tetragonal axis) at low temperatures.

[1] H. E. Bömmel, Phys. Rev. **96**, 220 (1954).
[2] See W. P. Mason's *Piezoelectric Crystals and Their Application to Ultrasonics* (D. Van Nostrand and Company, New York, 1950), p. 478.
[3] See G. Joos, *Theoretical Physics* (Hafner Publishing Company, New York, 1950), p. 562.
[4] See C. Kittel, *Introduction to Solid State Physics* (John Wiley and Company, New York, 1953), Chap. 12.
[5] See E. H. Sondheimer, Advances in Physics **1**, No. 1, 1–43 (1952).
[6] *International Critical Tables*. The value for tin is the value at room temperature for the tetragonal axis reduced by the factors found for polycrystal tin.

Can Helium-3 Be Expected to Exhibit Superfluidity at Sufficiently Low Temperatures?*

O. K. RICE
Department of Chemistry, University of North Carolina, Chapel Hill, North Carolina
(Received October 13, 1954)

LIQUID helium of atomic mass 4, below its λ point, is the only liquid which has so far been found to show superfluidity. According to the two-fluid hypothesis, it is actually only a certain part of the liquid helium, the superfluid, which has this property. The superfluid is a low-energy form of helium, there being a considerable energy gap between the level of an atom of superfluid and the lowest possible energy of an atom of normal fluid. Under these circumstances it is understandable that a force exerted at one point of a body of superfluid would be felt all through the superfluid, and that it would be set in motion almost immediately. For any force which tends to compress the superfluid, tends to raise the energy level of the entire mass. To avoid this,

the entire mass starts to move. And this kinetic energy is not easily transformed into molecular energy, since a considerable excitation is required to do this. It seems to me that this lack of transfer of kinetic energy into molecular energy is more easily understood if superfluid and normal fluid are separated in ordinary space as well as momentum space, as I have suggested in a number of papers.[1] But, in any event, it appears likely that the requirement for a superfluid is a continuous network of atoms in a low-energy level, which are not easily excited to a higher state, and which are still not rigidly fixed in their relative positions, as in a solid. These ideas are, of course, more or less familiar; we wish to discuss here how they can be applied to He^3, in the light of present knowledge.

In liquid He^3, there is no present evidence for the existence of a λ point. Furthermore, even at temperatures as low as 0.5°K there is an appreciable specific heat,[2] and magnetic measurements[3] indicate that antiparallel alignment of the nuclear spins does not occur to an appreciable extent until the temperature has been lowered to 0.2 or 0.3°K. I have discussed[4] this in terms of a model in which N He^3 atoms are considered to form $N/2$ pairs. At 0°K, all pairs are in the lowest rotational state and as the temperature is raised the nuclear magnetic susceptibility comes in through the pairs going to a higher rotational level. The excitation required is about 0.5 calorie per mole of He^3 atoms. This is low compared to the lowest excitation in liquid He^4, but, nevertheless, in the model I have considered the excitations are practically completely frozen out at 0.05°K. Thus it would appear that at this temperature the He^3 satisfies the conditions necessary for the appearance of superfluidity, since it is still liquid and there is a nonzero energy gap for excitation. Since there is no reason to expect a λ point, the superfluidity, if it occurs, should appear gradually as the temperature is lowered to a few hundredths of a degree. It certainly would be of interest to search for this effect at lower temperatures than have hitherto been used for investigations of this type on He^3.

On the other hand, if the helium behaves like an ideal Fermi-Dirac gas with particles of some effective mass different from that of the mass of an He^3 atom, then there is no such nonzero energy gap, and superfluidity, if it occurred at all, would be expected only exceedingly close to 0°K. The difference in the ideal Fermi-Dirac gas and the model I have mentioned above is well displayed in the difference in the magnetization curves at very low temperatures.[4] If the ideas presented here are correct, the Fermi-Dirac statistics in itself cannot prevent the appearance of superfluidity.

* Work supported by the U. S. Office of Naval Research.

[1] O. K. Rice, Phys. Rev. **76**, 1701 (1949); **93**, 1161 (1954).

[2] See, e.g., Osborne, Abraham, and Weinstock, Phys. Rev. **94**, 202 (1954).

[3] Fairbank, Ard, and Walters, Phys. Rev. **95**, 566 (1954).

[4] O. K. Rice, Phys. Rev. **97**, 263 (1955); P. J. Price, Phys. Rev. **97**, 259 (1955).

Ionization and Solubility in Semiconductors

HOWARD REISS AND C. S. FULLER
Bell Telephone Laboratories, Murray Hill, New Jersey
(Received December 2, 1954)

BY using the ideas of Reiss[1] concerning the effects of hole-electron equilibria on solubility, it can be shown that the concentration of donors, N_D, in a semiconductor (equilibrated with a constant activity solution) depends upon N_A, the concentration of acceptors, in the following manner.

$$N_D = N_A(N_D^0)^2/2K + [N_A^2(N_D^0)^4/4K^2 + (N_D^0)^2]^{\frac{1}{2}}, \quad (1)$$

in which N_D^0 is N_D for the case $N_A = 0$, and K is given by the relation:

$$N_D^0 = K/(K + n_i^2)^{\frac{1}{2}}, \quad (2)$$

where n_i is the intrinsic concentration of electrons for the temperature in question. It can be shown that K approximates the following dependence upon temperature:

$$\log K = A + B/T. \quad (3)$$

We have verified Eq. (1) within experimental error, using boron-doped silicon saturated with lithium by diffusion from a tin-lithium alloy phase. Table I compares observed and calculated values of N_D for various values of N_A at 350°, 440°, and 500°C. Values of n_i were obtained from the data of Morin and Maita,[2] while N_D^0 was simply measured and appears as the first figure (at each temperature) in the second column of Table I. All concentrations were determined by resistivity measurements together with appropriate values for hole electron mobilities.[3] For values of N_A in the neighborhood of 5×10^{18} cm^{-3}, Eq. (1) is an approximation (a more accurate expression is available but it requires the use of additional parameters) and predicts a high result, the discrepancy being more pronounced at lower temperatures.

TABLE I. Observed and calculated values of concentration of donors, N_D, for various values of concentration of acceptors, N_A.

N_A(cm^{-3})	Obs N_D(cm^{-3})	Calc N_D(cm^{-3})
350°C		
1.1×10^{15}	1.12×10^{16}	
1.78×10^{16}	2.01×10^{16}	1.97×10^{16}
1.34×10^{17}	8.6×10^{16}	1.05×10^{17}
5.63×10^{18}	3.25×10^{18}	4.36×10^{18}
440°C		
1×10^{15}	6.55×10^{16}	
1.78×10^{16}	7.17×10^{16}	7.23×10^{16}
1.47×10^{17}	1.20×10^{17}	1.45×10^{17}
5.84×10^{18}	1.85×10^{18}	4.3×10^{18}
500°C		
1.1×10^{15}	1.08×10^{17}	
1.73×10^{16}	1.30×10^{17}	1.13×10^{17}
1.34×10^{17}	1.46×10^{17}	1.6×10^{17}
5.59×10^{18}	2.95×10^{18}	3.59×10^{18}

By using Eqs. (2) and (3), it is possible to calculate, in part, the solubility-temperature curve (solubility $= N_D^0$) for pure lithium in silicon. The measured[4] points appear in Fig. 1, and the full line represents the probable

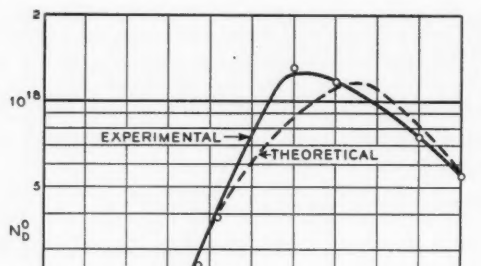


FIG. 1. Comparison of calculated and experimental solubility-temperature curves for Li in Si. The experimental results are replotted from reference 4 neglecting the result at 500°C which appears to be in error.

smooth experimental curve. At low temperatures (below 450°C), $K \gg n_i^2$ so that, according to (2), $K = (N_D^0)^2$. Employing the experimental data for N_D^0 below 450°C, it is thus possible to determine A and B, and, therefore, K at any temperature. The result of this calculation appears as the dashed line in Fig. 1. The occurrence of the maximum is here attributed to the increase in concentration of intrinsic electrons suppressing the ionization of the donor and thus lowering the solubility. A more detailed paper will be published later. The authors would like to thank J. A. Ditzenberger and A. J. Pietruszkiewicz for assistance with the experiments.

[1] Howard Reiss, J. Chem. Phys. **21**, 1209 (1953).
[2] F. J. Morin and J. P. Maita, Phys. Rev. **96**, 28 (1954).
[3] M. B. Prince, Phys. Rev. **93**, 1204 (1954).
[4] C. S. Fuller and J. A. Ditzenberger, Phys. Rev. **91**, 193 (1953).

K Band in Additively Colored KCl at 18°K

FELIX E. GEIGER, JR.
Naval Research Laboratory, Washington 25, D. C.
(Received December 2, 1954)

THE *K* band was first observed by Kleinschrod[1] in additively colored KCl. However, the absorption curves of the crystal were measured at room temperature with consequent poor resolution between the *F* and *K* band. The theory of Mott and Gurney[2] as well as the recent absorption curves[3–5] of x-rayed alkali halide crystals gave new impetus to work on *K* centers.[6] These centers have also been designated by the letter V_0 because of their apparent *V*-band type behavior. Although the x-ray data show clearly the *K* band, especially in KCl, K centers in additively colored crystals seem less easily obtainable. Pick's[7] curves even at 38°K show a complete absence of the *K* band, and so do the recent curves of Petroff[8] at 100°K. The technique in growing crystals and coloring them used by Kleinschrod, Pick, and Petroff was presumably the same. Equally negative results are evident in the absorption curves of Scott and Bupp[9] and the evaporated KCl of Kaiser.[10,11] There are no published low-temperature measurements of *F* and *K* bands in additively colored crystals, which allow of direct comparison with the x-ray data of Duerig and Markham.[4]

It has been possible to obtain a clearly defined *K* band at approximately liquid helium temperature. The techniques of temperature measurement at liquid helium temperature, and the low-temperature absorption cell are described in a forthcoming paper in the *Review of Scientific Instruments*. Figure 1 shows the absorption

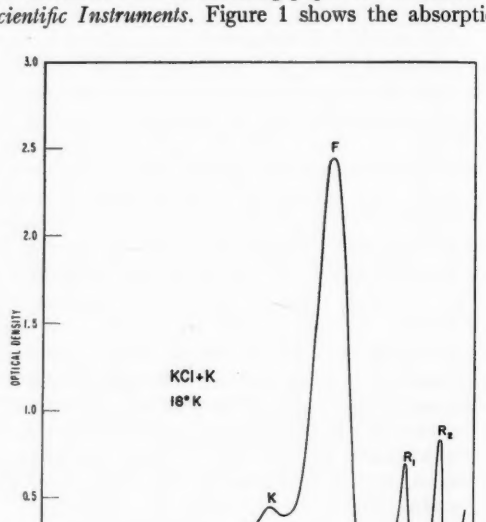


FIG. 1. Absorption curve for additively colored KCl at 18°K.

spectrum of additively colored KCl at 18°K. The crystals, obtained from Harshaw Chemical Company, were colored by heating in potassium at 477°C for 28 hours, then quenched in icewater. The absorption measurements were made with a Cary recording spectrophotometer. Table I summarizes the wavelength measurements on *K*, *F*, R_1, and R_2 bands at 18°K and 83°K. The wavelength determinations were checked and corrected against the spectrum of a mercury arc lamp. The

TABLE I. Wavelength of the maximum of the F, K, and R bands in additively colored KCl.[a]

Temperature °K	K	F	R_1	R_2
18	464 mμ 2.68 ev	533 mμ 2.33 ev	733 mμ 1.70 ev	656 mμ 1.89 ev
83	468 mμ 2.65 ev	537 mμ 2.31 ev	732 mμ 1.70 ev	658 mμ 1.89 ev

[a] Thickness of crystal 0.016 inch.

broadness of the bands may add an uncertainty of approximately 2 mμ, and the measurements certainly cannot be made closer than to 2 or 3 mμ, giving a total uncertainty in the wavelength of roughly 4 mμ. The $K(V_0)$ band measured by Duerig and Markham has a wavelength of 455 mμ or 2.73 ev at a measurement temperature of 5°K.

The ratio of the peak intensity of the F and K bands and its possible variation is also of considerable interest. Table II summarizes the results on these ratios. The

TABLE II. Densities and ratios of densities of F, K, and R bands in KCl measured at 100°K.

Crystal No.	Thickness inch	Optical density F	K	R_1	R_2	F/K	F/R_1	F/R_2	Remarks
1	0.016	2.61	0.44	0.62	0.74	5.9	4.2	3.5	K band resolved
6	0.009	2.57	0.38	0.43	0.43	6.8	6.0	6.0	K band with very
2	0.013	2.04	0.38	0.48	0.55	5.4	4.3	3.7	slight minimum be-
10	0.021	1.98	0.36	0.38	0.43	5.5	5.2	4.6	tween K and F band

crystals were colored and quenched as indicated above, and no subsequent bleaching, heating, or quenching was performed on them. The F/K ratios differ significantly from the x-ray data of Dorendorf and Markham. Seitz[12] gives 21 for this ratio, based on Duerig and Markham's data; Dorendorf's curves give a figure of roughly 13.

I should like to thank Dr. H. Friedman for suggesting a renewed investigation of additively colored alkali halides, and Dr. D. Dexter for several discussions on K centers.

[1] F. G. Kleinschrod, Ann. Physik **27**, 97 (1936).
[2] N. F. Mott and R. W. Gurney, *Electronic Processes in Ionic Crystals* (Oxford University Press, London, 1940), p. 114.
[3] H. Dorendorf, Z. Physik **129**, 166 (1950).
[4] W. H. Duerig and J. J. Markham, Phys. Rev. **88**, 1043 (1952).
[5] Casler, Pringsheim, and Yuster, J. Chem. Phys. **18**, 887 (1950).
[6] H. W. Etzel and F. E. Geiger, Jr., Phys. Rev. **96**, 225 (1954).
[7] H. Pick, Ann. Physik **31**, 365 (1938).
[8] S. Petroff, Z. Physik **127**, 443 (1950).
[9] A. B. Scott and L. P. Bupp, Phys. Rev. **79**, 341 (1950).
[10] R. Kaiser, Z. Physik **132**, 482 (1952).
[11] Both Pick and Petroff heated and then quenched their crystals before measurement, Kleinschrod apparently did not, although this is not explicitly mentioned. Scott and Bupp show F, R_1, R_2, and M bands at 80°K but no K band for a crystal which had not been heated and quenched before the absorption measurement. Kaiser's curves of evaporated KCl taken at 90°K show R and M bands but apparently no K bands.
[12] F. Seitz, Revs. Modern Phys. **26**, 7 (1954).

Upper Limit for the Lifetimes of Excited States of Ni^{60}†

Z. BAY, V. P. HENRI, AND F. MCLERNON
The George Washington University, Washington, D. C.
(Received October 25, 1954)

AN upper limit of 10^{-11} second for the lifetimes of the two excited states of Ni^{60} following the beta decay of Co^{60} has been determined. The E_2 transition energies are 1.17 and 1.33 Mev.

A thin Co^{60} source (~20 microcuries) electrolytically deposited on a Cu disk (440 mg/cm²) is placed between two diphenyl acetylene crystals, called A (8×8×20 mm³) and B (1.5×8×8 mm³), each facing a 1P21 photomultiplier. Between the Co^{60} source and crystal B an Al absorber (98 mg/cm²) can be interposed, preventing any betas from reaching B; betas can never reach crystal A.

With the Al absorber alternatively in and out of place, we measured a $[\gamma_1\gamma_2+\gamma_2\gamma_1]$ coincidence curve, called $[\gamma\gamma]$, and a composite $[\beta\gamma_1+\beta\gamma_2]+[\gamma_1\gamma_2+\gamma_2\gamma_1]$ coincidence curve, called $[\beta\gamma]+[\gamma\gamma]$. The difference between these two curves is the $[\beta\gamma]$ curve. No "prompt" source is needed in this experiment because application of moment theorems[1] to the $[\beta\gamma]$ and $[\gamma\gamma]$ curves yields for the mean lives of the upper (θ_1) and of the lower (θ_2) excited states:

$$\theta_1=(5/4)\Delta\mu_1-\tfrac{1}{4}\{9(\Delta\mu_1)^2-4(\Delta\mu_2-2\mu_1[\gamma\gamma]\Delta\mu_1)\}^{\frac{1}{2}}, \quad (1)$$

and

$$\theta_2=-\tfrac{1}{2}\Delta\mu_1+\tfrac{1}{2}\{9(\Delta\mu_1)^2-4(\Delta\mu_2-2\mu_1[\gamma\gamma]\Delta\mu_1)\}^{\frac{1}{2}}, \quad (2)$$

where $\Delta\mu_n$ is the difference between the nth normalized moments of the $[\beta\gamma]$ and $[\gamma\gamma]$ curve and $\mu_1[\gamma\gamma]$ is the normalized first moment (centroid) of the $[\gamma\gamma]$ curve.

In our coincidence equipment, shown in Fig. 1, the outputs of detectors A and B are branched to the two coincidence circuits C and D. C is a diode bridge coincidence circuit[2] with no variable time delay in its channels. Its resolving time (5×10^{-9} second) is chosen to exceed the two expected lifetimes and all time lags present by at least one order of magnitude. C yields the total coincidence counting rate N_0.[3] D is a differential

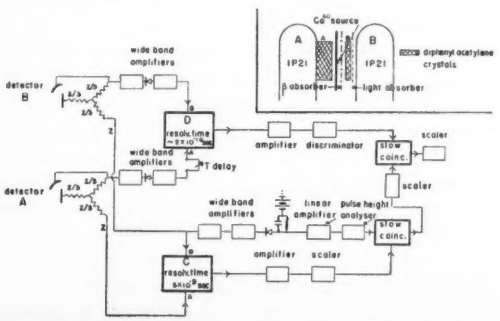


FIG. 1. Schematic diagram of circuits for detecting coincidences.

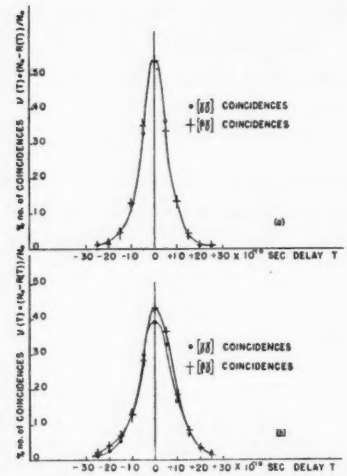


FIG. 2. $[\gamma\gamma]$ and $[\beta\gamma]$ delayed coincidence curves. In Fig. 2a, $[\gamma\gamma]$ and $[\beta\gamma]$ coincidences were selected by use of a pulse height analyzer as described in the text. In Fig. 2b, $[\gamma\gamma]$ coincidences were obtained while using a light absorber between crystal B and its photomultiplier.

coincidence circuit[4–6] with a variable time delay T in its channel A. For coincidences, D gives positive outputs [counting rate $D_+(T)$] if the pulse in channel A arrives later than that in B, negative outputs [counting rate $D_-(T)$] for the opposite case, and no output for the most accurate coincidences occurring. Hewlett Packard 460A amplifiers amplify the D input pulses to be certain that, whenever C detects a coincidence, D has a response. The operation of D is independent of pulse lengths and its resolving time can be as low as 10^{-10} second. Slow coincidences (10^{-5} second) between the outputs of C and D produce a counting rate $R(T)$. Plotting $\nu(T)=[N_0-R(T)]/N_0$ *versus* T, we obtain the "reduced coincidence curves."

From $\nu(T)$ we compute $\bar{t}$, a mean of the random time lags occurring in the equipment.[3,5] A small value of $\bar{t}$ is important for good statistical accuracy of time measurements.[3,7] When the photomultiplier pulses were used without any pulse-shaping, $\bar{t}$ was 8×10^{-10} second. This has been reduced to 4.5×10^{-10} second by a clipping method (anode pulses shortened by dynode pulses).[8] The observed $\bar{t}$ is not far from the theoretical lower limit for the radiations and the detectors used.[9]

In our case, the usual requirement for using the same amplitude distributions to measure two different coincidence curves applies only to channel B, since there is no change in the excitation of channel A for the measurement of the two curves. In order to obtain a pulse height selection, Channel B is branched at the input to C and the branched pulses sent through a diode, followed by an RC circuit (10^{-6} second), to a linear amplifier and then to an Atomic Instrument Company 501 pulse height analyzer. In order not to overload the input of the pulse height analyser, a preselection was performed by biasing the diode. The selected pulses were used to gate (5×10^{-6} second) the outputs of C. About 15 percent of the $[\gamma\gamma]$ and 5 percent of the $[\beta\gamma]$ coincidences were selected.

The curves in Fig. 2a show no difference exceeding statistical fluctuations, thereby indicating very short lifetimes. Therefore, Eqs. (1) and (2) can only be used to set an upper limit for θ_1 and θ_2 giving $\theta_1+\theta_2/2=\Delta\mu_1$ (shift of the centroids).

The curves in Fig. 2a represent one set of measurements corresponding to $\sim$4000 coincidences for each curve. Three such sets gave for the average displacement of the centroids 0.1×10^{-12} second with a maximum deviation of 2.4×10^{-12} second. The calculated statistical error for the entire set of 12 000 coincidences is $\sim1\times10^{-11}$ second. The calculated average difference of times of flight between β's and γ's reaching crystal B (distance from the source = 1.5 mm, crystal thickness = 1.5 mm) is $\sim10^{-12}$ second. Errors due to coincidences resulting from backscattering have been checked and were also found to be negligible. To check our technique for very short time measurements, time-of-flight experiments were performed by varying the position of crystal A. Times of flight as short as a few times 10^{-11} second have been measured and were found to be within the above statistical error, as expected. We consider therefore 10^{-11} second to be the upper limit for the mean lives of both excited states of Ni^{60}. This value is compatible with theoretical estimates.[10,11] (Earlier experiments gave upper limits approximately two orders of magnitude higher.)[12,13]

Next, using Co^{60} as a "prompt" source, we tried to work out a method for similar measurements avoiding use of a pulse height analyser which causes $\sim$90 percent loss of the useful coincidences. We tried to equalize the amplitude distributions for the two radiations by using a light absorber between the crystal and the tube which detect the higher energies of excitation. The decay time of the scintillator is widely independent of energy and (of course) of the number of photons utilized.

First, we determined how the time measurements are influenced by the different β and γ amplitude distributions. Curves 1 and 2 in Fig. 3 show the broad amplitude distributions of β and γ singles in channel B measured in coincidence with the C outputs. The ratio of the average

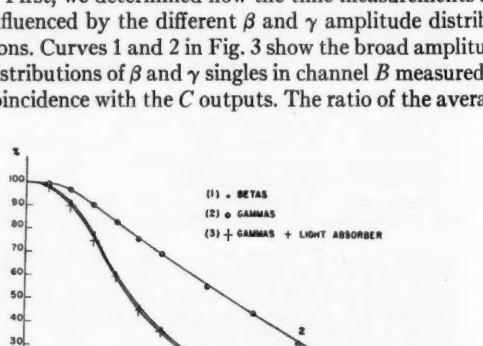


FIG. 3. Integral amplitude distributions of singles in channel B, measured in coincidence with the C outputs.

amplitudes $\bar{A}_\gamma/\bar{A}_\beta$ is 1.6. The shift of the centroids of the corresponding delay curves, taken without pulse height selection, was $(5\pm1)\times10^{-11}$ second.

Using next a light absorber (wire mesh) of the proper transparency between crystal B and its photomultiplier when measuring γ's, it was possible to adjust the γ amplitude distribution to that of the β's so that the rms of the differences between corresponding ordinates was 4.9 percent (curve 3, Fig. 3). The shift of the centroids of the corresponding delay curves (Fig. 2b) was $(0.5\pm1)\times10^{-11}$ second, i.e., within statistical accuracy. From this it appears that the equalization of amplitude distributions to within a few percent, permits one to make time measurements by the coincidence method with an accuracy of $\sim10^{-11}$ second.

† This work was supported by the joint program of the U. S. Office of Naval Research and the U. S. Atomic Energy Commission.

[1] Z. Bay, Phys. Rev. **77**, 419 (1950).
[2] Z. Bay, Rev. Sci. Instr. **22**, 397 (1951).
[3] Bay, Kanner, and Henri, Oak Ridge National Laboratory Technical Report ONR Contract 168(00), September, 1953; and to be published.
[4] Z. Bay, Phys. Rev. **83**, 242 (1951).
[5] Bay, Cleland, and McLernon, Phys. Rev. **87**, 901 (1952).
[6] Bay, Henri, and McLernon (to be published).
[7] Kanner, Bay, and Henri, Phys. Rev. **90**, 371 (1953).
[8] Bay, Henri, and McLernon, Phys. Rev. **90**, 371 (1953).
[9] R. F. Post and L. I. Schiff, Phys. Rev. **80**, 1113 (1950).
[10] V. F. Weisskopf, Phys. Rev. **83**, 1073 (1951).
[11] Aeppli, Frauenfelder, Heer, and Rüetschi, Phys. Rev. **87**, 379 (1952).
[12] M. Deutsch and W. E. Wright, Phys. Rev. **77**, 139 (1950).
[13] Gorodetzki, Knipper, Armbruster, and Gallmann, J. phys. et radium **14**, 550 (1953).

Inelastic Collision Cross Sections at 1.0-, 4.0-, and 4.5-Mev Neutron Energies

J. R. Beyster, R. L. Henkel, and R. A. Nobles
Los Alamos Scientific Laboratory, Los Alamos, New Mexico
(Received November 22, 1954)

In recent years, sphere transmission measurements have been used successfully to determine inelastic collision cross sections.[1–5] In this method the transmission of a spherical shell is measured by surrounding either the neutron detector or the neutron source. To minimize the difficulties caused by the highly anisotropic $T(p,n)He^3$ neutron source used in this experiment, the neutron detector was surrounded instead of the source. The maximum spread in neutron energies was 160 kev for the 1-Mev source and about 70 kev for the 4.0 and 4.5-Mev sources. A high-pressure hydrogen gas proportional counter detected the 1-Mev neutrons. At 4.0 and 4.5 Mev, the neutron detector was a hydrogen recoil scintillation counter consisting of nine spheres of plastic phosphor, 0.100 inch in diameter, separated by quartz to reduce gamma-ray sensitivity.[6] Each detector was positioned at about 30 inches from the tritium gas target in the forward direction. It was operated as a biased neutron detector with about ten energy thresholds allowing ten transmissions to be measured simultaneously. This procedure was adopted to obtain information concerning the energy spectrum of inelastic neutrons produced in the spheres.

When calculating inelastic cross sections from observed sphere transmissions, it is important to consider the probability of multiple neutron scattering. The errors which are made by not considering multiple scattering were not investigated accurately until a few years ago. At that time, Bethe[2–4] proposed a simple and accurate method of evaluating multiple scattering in a spherical geometry which agreed with both experimental observation and detailed Monte Carlo calculations of this effect.

We included the following effects in the numerical analysis of the experimental data: (1) the multiple scattering in the sphere, (2) the relative variation in energy and intensity of neutrons over the solid angle subtended by the sphere, (3) the effect of loss of energy in elastic collisions with resulting reduction in counting sensitivity, and (4) the effect of finite detector size inside the sphere. Corrections due to these effects can be computed using basically the Bethe method of analysis. Since the magnitude of the third correction is large for light elements and the higher energy thresholds, this effect was computed by the Monte Carlo method for increased accuracy. All corrections listed above depend to varying extents on the shape of the elastic scattering angular distributions and on the total neutron cross sections. The known angular distributions at 1 Mev[7] and the recently measured distributions at 4.1 Mev[8] were used. To speed up the numerical analysis the entire computing problem was coded for the Los Alamos Maniac Computer by E. D. Cashwell, C. J. Everett, and J. M. Kister.

The cross sections given in Table I for 1-Mev neutron energy were taken with the proportional counter biased at 750 kev. With the exception of gold and silver, the cross sections are relatively insensitive to the energy threshold of the detector and are probably the entire inelastic collision cross sections. This cross section for

Table I. Inelastic collision cross sections.

Element	1 Mev	4.0 Mev	4.5 Mev
Beryllium		0.62±0.05	
Carbon	−0.09±0.14	0.04±0.08	
Aluminum	0.04±0.08	0.75±0.05	0.72±0.06
Titanium		1.28±0.09	1.16±0.09
Iron	0.41±0.04	1.42±0.07	1.34±0.07
Nickel		1.35±0.09	1.50±0.09
Copper	0.21±0.05	1.60±0.07	1.60±0.06
Zinc	0.10±0.06	1.69±0.06	1.81±0.09
Zirconium		1.56±0.07	1.59±0.09
Silver	1.61±0.16	2.05±0.10	2.02±0.11
Cadmium	0.99±0.06	2.05±0.10	2.12±0.11
Tin	0.07±0.05	2.09±0.10	2.18±0.10
Tungsten		2.58±0.20	2.56±0.20
Gold	1.63±0.10	2.75±0.12	2.69±0.13
Lead	0.23±0.04	1.84±0.08	2.02±0.16
Bismuth	0.12±0.04	1.98±0.10	2.19±0.10

gold at 1 Mev is estimated to be about 1.8 barns. The 1-Mev data were compared with the inelastic cross sections obtained by Walt and Barschall.[7] For the majority of the elements, the agreement is within the experimental errors.

The cross sections given for 4.0- and 4.5-Mev neutron energies were obtained with the scintillation detector biased at 3.4 and 3.8 Mev respectively. The values do not change rapidly with the energy threshold of the detector. These data agree very well with inelastic collision cross sections recently determined at 4.1-Mev neutron energy in the elastic scattering angular distribution experiments.[8]

A more extensive report on the experimental procedure, numerical analysis, and results is being prepared.

[1] H. H. Barschall, Revs. Modern. Phys. **24**, 120 (1952).
[2] H. A. Bethe, Los Alamos Report LA-1428, 1952 (unpublished).
[3] Bethe, Beyster, and Carter, Los Alamos Report LA-1429 (unpublished).
[4] H. A. Bethe, Report of Second Conference on Medium Energy Nuclear Physics at University of Pittsburgh, 1953 (unpublished).
[5] Phillips, Davis, and Graves, Phys. Rev. **88**, 600 (1952).
[6] McCrary, Taylor, and Bonner, Phys. Rev. **94**, 808(A) (1954).
[7] M. Walt and H. H. Barschall, Phys. Rev. **93**, 1062 (1954).
[8] M. Walt and J. R. Beyster, Bull. Am. Phys. Soc. **29**, No. 8, 31 (1954).

Multiplicity of Neutrons from the Spontaneous Fission of Californium-252*

DONALD A. HICKS, JOHN ISE, JR., AND ROBERT V. PYLE
Radiation Laboratory, Department of Physics, University of California, Berkeley, California
(Received November 22, 1954)

THE neutron number distributions from the spontaneous fission of some of the transuranic elements are being measured with a neutron detector of high efficiency. The first data to be analyzed are from a sample of Cf^{252} with about 300 spontaneous fissions per minute.

The material was mounted in a fission chamber placed at the center of a cylindrical tank of cadmium-loaded liquid scintillator. The dimensions and construction of this detector are nearly identical with the design of the Los Alamos group.[1] A pulse from the fission chamber triggered the sweep of an oscilloscope. The fission neutrons were moderated in the toluene and captured in the cadmium with a mean lifetime of 20 microseconds, and some of the resulting γ rays produced pulses in the scintillator which were displayed on the scope trace and photographically recorded.

The numbers of fissions with 0, 1, 2, etc., *detected* neutrons are given in Table I. A correction has been made for a background of 0.75 percent. The average number of neutrons per spontaneous fission of this nucleus has been measured by Crane *et al.*[2] to be 3.10 ± 0.18.[3] The 8494 fissions reported here gave an average of 1.43 neutrons, from which we conclude that our detection efficiency was 46.0 ± 2.8 percent (probable error) for this measurement. A Monte Carlo calculation made at the Livermore laboratory indicates that this efficiency is constant over the energy interval in which nearly all the fission neutrons are expected to lie. Using this value of the efficiency we calculated the number of fissions *vs* the *true* numbers of neutrons per fission; the results are given in Table I.

It is seen that there is a large probability that all Cf^{252} fissions emit at least one neutron, and the fraction of spontaneous fissions giving seven or more neutrons is certainly less than one percent.

The multiplicities of neutron production by spontaneous fission of natural uranium have been measured by Geiger and Rose,[4] using equipment with a neutron detection efficiency of three percent.

They were able to fit their data to a Poisson distribution, although other forms were not excluded. It can be shown that if the true distribution is binomial with a mean number of neutrons m, and the efficiency with which a single neutron is detected is e, then the observed

TABLE I. Observed and calculated neutron multiplicities from the spontaneous fission of Cf^{252}.

No. of neutrons	0	1	2	3	4	5	6	7
Observed distribution	1596	3200	2445	1001	214	33	5	0
True distribution	−65 ±100	+999 ±340	+1407 ±770	+2970 ±950	+2763 ±1150	−107 ±600	+528 ±250	0

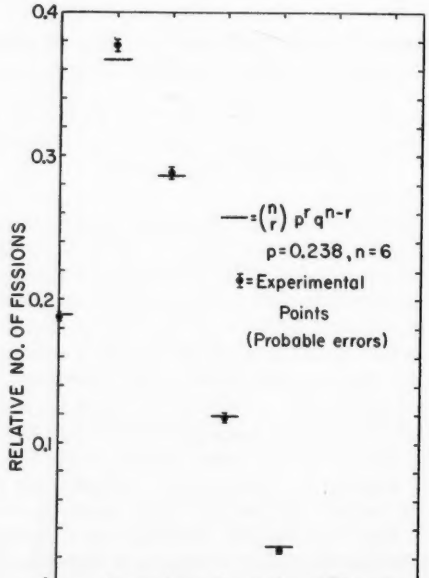


FIG. 1. Observed neutron number distribution arising from the spontaneous fission of Cf^{252}.

distribution will also be of the binomial form $\binom{n}{r}p^r q^{n-r}$, with $p=em/n$. If we take $n=6$, then $p=0.238$. It is seen from Fig. 1 that the experimental points agree well with this description. The observed points will also fit the expansion with $n=7$. Conversely, if the observed distribution is binomial, the true distribution is also binomial. Because of the magnitude and probable error of the detection efficiency, we are not able to state that the true distribution is binomial.

We are in the process of improving the efficiency of the detector as well as the statistics of the experiment. A more detailed discussion of the method and results will be submitted shortly, along with data from other elements and isotopes.

The use of a large liquid scintillator tank as an efficient neutron moderator and detector was suggested to us by Walter E. Crandall. Discussions with George P. Millburn and R. L. Gluckstern concerning treatment of the data were very helpful. We are grateful to Stanley Thompson and Albert Ghiorso of this laboratory for supplying the sample, and wish to thank Edward Leshan for the Monte Carlo calculations, Frank Adelman for much preliminary work, and Stephen Kahn and Edith Goodwin for much of the processing and reading of the film. This work was done under the supervision of C. M. Van Atta.

* This work was done under the auspices of the U. S. Atomic Energy Commission.

[1] Cowan, Reines, Harrison, Anderson, and Hayes, Phys. Rev. 90, 493 (1953).

[2] Crane, Higgins, and Thompson, Phys. Rev. 97, 242 (1955).

[3] The probable error quoted here differs slightly from that in the Letter by Crane *et al.* because of further information from Mound Laboratory concerning the neutron standard used.

[4] K. W. Geiger and D. C. Rose, Can. J. Phys. 32, 498 (1954).

Evidence for a Second Naturally Occurring Isotope in Tantalum*

J. E. Evans, E. G. Joki, and R. R. Smith

Phillips Petroleum Company, Atomic Energy Division, Idaho Falls, Idaho

(Received November 24, 1954)

A SMALL "resonance shaped" peak was observed at 0.433 ± 0.004 ev during high-resolution tantalum total neutron cross section measurements with the Materials Testing Reactor crystal spectrometer. The data are presented in Fig. 1. Evidence for this peak is also to be found in Christensen's data.[1] The 240 planes of an NaCl monochromating crystal were used. The width at half-maximum of the spectrometer resolution function at 0.43 ev was 0.007 ev. A tantalum metal sample 0.505 ± 0.001 cm thick was used. It was obtained from the Fansteel Metal Corporation, and had a stated purity greater than 99.9 percent. The measured density of the sample was 16.58 ± 0.02 g/cm^2.

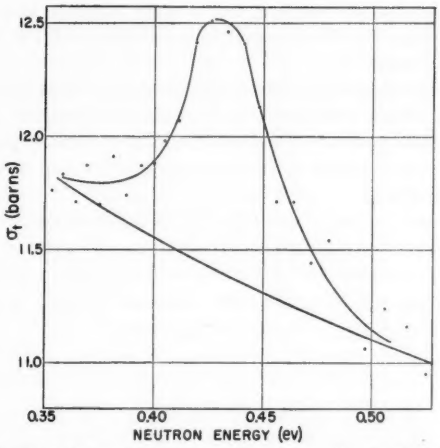


FIG. 1. The total neutron cross section of tantalum in the 0.43-ev region. The lower curve is a $1/v$ fit to the nonresonance data.

The resonance effect was isolated by plotting cross section values against $E^{-\frac{1}{2}}$ and subtracting values given by a visual straight-line fit of the nonresonance data from the experimental values. The resulting data are plotted in Fig. 2. The observed width of the peak at half-maximum is 0.048 ± 0.005 ev and the observed height is 1.16 ± 0.08 barns. The Breit-Wigner single-level formula, corrected for Doppler broadening by Sailor's variation of the trial and error method,[2] was used to obtain the solid curve in Fig. 2. After making a small instrument resolution correction (0.001 ev) on the width, the resonance parameters for the element are: $E_0=0.433\pm0.004$ ev, $\sigma_0=1.59\pm0.11$ barns, $\Gamma=0.030\pm0.005$ ev, and $g\Gamma_n=0.8\times10^{-3}$ ev. The location of the resonance is outside the recommended range of applica-

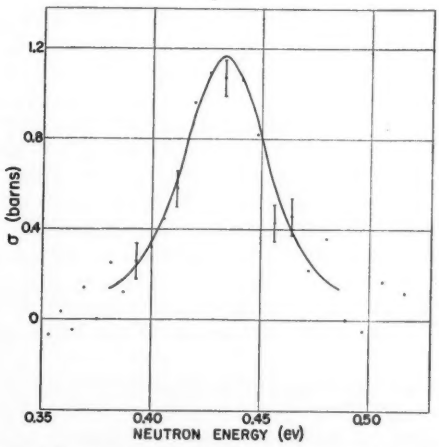


FIG. 2. A linear plot of the tantalum resonance data. The $1/v$ component has been subtracted from the total cross section data of Fig. 1. The solid curve is a calculated single level Breit-Wigner curve corrected for Doppler broadening. The parameters used are those given in the text for the element with natural isotopic abundance. The standard deviations indicated are those due to counting statistics.

tion for Sailor's method, but a more accurate Doppler correction would reduce Γ below 0.030 ev by a very small amount.[3]

The energy of the peak is quite high for the peak to be a scattering cross section discontinuity caused by a crystal edge in the sample. In addition, such an effect would result in a sawtooth peak rather than a symmetrical one.

A search of existing cross-section data[4] indicates that the 0.43-ev peak is probably not due to a contaminant. If it is assumed that any element having such a narrow resonance (0.030 ev) must have an atomic weight $A>150$,[5] most of the remaining elements can be ruled out by their cross section data. Gd, Tb, and the radioactive elements Ra, Ac, and Pa are the only elements for which measurements for the 0.43-ev region have not been published. The radioactive elements can be ruled out as possible contaminants since the sample shows no alpha activity. Effects of a high resonance at 0.43 ev in Gd would be expected to extend into regions where measurements exist, even though there is a gap in existing data from 0.2 to 1 ev. If the resonance is assumed to be caused by a Tb impurity, and its width and atomic weight are compared with the data from Fig. 1 of the paper of Hughes and Harvey,[5] then the point for this resonance is seen to be farther below the solid curve than any other point plotted there. Thus a Tb impurity is probably not the cause of the 0.43-ev peak. If the resonance is due to Ta^{181}, the value of $g\Gamma_n$ would be expected to be of the order of 10^{-4} ev instead of 0.8×10^{-8} ev.

The above considerations led to the hypothesis that another stable (or long-lived) isotope of tantalum exists. Earlier measurements[6] placed the upper abundance limit of a second isotope at 100 ppm (parts per million). If such an isotope exists and has a resonance at 0.43 ev, it would be expected to have a peak height of the order of 10^4 to 10^5 barns. Considering the observed peak, this would correspond to a relative isotopic abundance of the order of 20 to 200 ppm. The other isotope would probably be within two mass numbers of 181 since there are no known cases of odd-Z elements with $Z>19$ having more than two naturally occurring isotopes. When such elements do have more than one isotope, not more than two mass numbers separate them. Thus the atomic weight of a second Ta isotope should be 181 ± 2. It is quite probable that the second isotope is radioactive since the second naturally occurring isotopes of other odd-Z elements with Z between 64 and 76 are radioactive.

The authors suggested the hypothesis, that a second isotope of Ta exists, as one which could be proved or disproved by use of the recently completed high intensity mass spectrograph at Knolls Atomic Power Laboratory, Schenectady, New York. The KAPL group accepted the suggestion and obtained the results given in an accompanying "Letter to the Editor."[7] If one uses their measured value of 123 ± 3 ppm for the relative abundance of Ta^{180}, the present total cross section data give the following isotopic single-level Breit-Wigner parameters for the 0.43-ev neutron resonance in Ta^{180}: $E_0=0.433\pm0.004$ ev, $\sigma_0=12\,900\pm1000$ barns, $\Gamma=0.030\pm0.005$ ev, and $g\Gamma_n=0.6\times10^{-4}$ ev. These parameters are especially interesting since Lu^{176} is the only other odd-odd isotope with $Z>7$ for which resonance parameters have been determined.[8] The assignment of this isotope to an odd-proton–odd-neutron position strengthens the possibility that it is radioactive since all other known odd-odd nuclei with $Z>7$, except V^{50}, are radioactive.

* Work performed under the auspices of the U. S. Atomic Energy Commission.

[1] R. L. Christensen, Phys. Rev. **92**, 1509 (1953).
[2] V. L. Sailor, Phys. Rev. **91**, 53 (1953).
[3] M. W. Holm (private communication).
[4] *Neutron Cross Sections*, U. S. Atomic Energy Commission Report AECU-2040 (Office of Technical Services, Department of Commerce, Washington, D. C., 1952); Supplements to AECU-2040.
[5] D. J. Hughes and J. A. Harvey, Nature **173**, 942 (1954).
[6] J. R. White and A. E. Cameron, Phys. Rev. **74**, 991 (1948).
[7] White, Collins, and Rourke, accompanying Letter [Phys. Rev. **97**, 566 (1955)]. We are grateful to the Knolls Atomic Power Laboratory group for communicating their results to us prior to publication.
[8] Foote, Landon, and Sailor, Phys. Rev. **92**, 656 (1953).

New Naturally Occurring Isotope of Tantalum

F. A. WHITE, T. L. COLLINS, JR., AND F. M. ROURKE
Knolls Atomic Power Laboratory, General Electric Company, Schenectady, New York*

(Received November 17, 1954)

WE have carefully examined the mass spectrum in the region of Ta^{181} and find that Ta^{180} exists as a naturally occurring isotope of low abundance.[1] New isotopic values for tantalum are as follows:

Isotope	Abundance (%)	
180	0.0123 ± 0.0003	
181	99.988	
177, 178, 179	<0.0003	not observed
182, 183	<0.0002	

The reason for this isotope not being previously observed[2] is probably, in part, that it is obscured by scattered Ta^{181} ions when the beam intensity is reasonably large. In the present measurement, a tantalum ribbon filament was used as a thermal ionization source. A twenty-stage electron multiplier was used to count individual positive ions[3] and the 180/181 ratio was actually determined by evaluating a counting ratio rather than by interpreting a rate meter integral curve, (see Fig. 1). The Ta^{180} peak corresponded to an ion current of about 10^{-18} ampere.

Two different mass spectrometers were used to make isotopic identification of Ta^{180}. The first was a single 6-inch radius-of-curvature analyzer of 60°. The second was a two-stage analyzer having 12-inch radius −90°

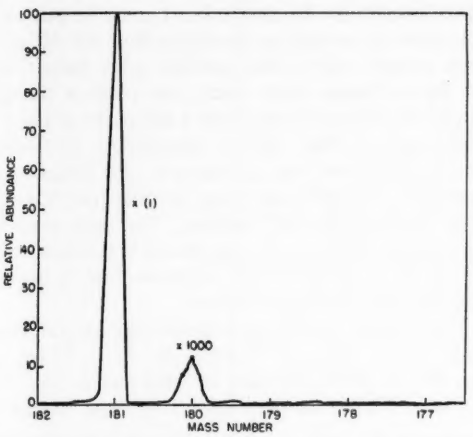


FIG. 1. Tantalum mass spectrum obtained from a two-stage magnetic analyzer utilizing two 90°, 12-inch radius-of-curvature sectors in tandem. The curve is a trace of the output of a twenty-stage electron multiplier detector when integrated by a counting rate meter.

magnetic sectors. It was this latter instrument[4] which was specifically designed for high-abundance sensitivity work, in which the above measurements were made. The 6-inch radius instrument yielded the same values, but the poor resolution did not allow for as much precision in the 180/181 ratio. The absence of isobars of tungsten and hafnium is indicated and oxides of lighter element parents are also excluded (Fig. 1). In addition, the mass difference between 180 and 181 was measured to be 1 mass unit within 1 percent, thus excluding the 180 peak as arising from any simple compounds. Further, the 180/181 positive ion ratio was found to be constant over a wide temperature range for many samples.

* Operated by the General Electric Company for the U. S. Atomic Energy Commission.

[1] The existence of another isotope of tantalum was suggested as a possibility in order to explain a small neutron cross-section resonance at 0.44 ev, not thought to be ascribable to the known nucleus, Ta^{181}, J. E. Evans (private communication).

[2] J. R. White and A. E. Cameron, Phys. Rev. 74, 991 (1948).

[3] F. A. White and T. L. Collins, Applied Spectroscopy 8, 17 (February, 1954).

[4] F. A. White and T. L. Collins, Applied Spectroscopy 8 (November, 1954).

Proton-Gamma Resonances in Magnesium

S. E. HUNT AND D. A. HANCOCK

Research Laboratory, Associated Electrical Industries Ltd., Aldermaston, Berkshire, England

(Received November 26, 1954)

IN a recent letter on the ground state of Al^{26}, Kluyver *et al.*,[1] stated that resonances occurring at 315, 387, and 436 kev had been interpreted as $Mg^{26}(p,\gamma)Al^{27}$ resonances by Hunt and Jones.[2]

We feel that the summary of the paper referred to by Kluyver may have led to this misunderstanding, since in the work reported no attempt was made to check Tangen's initial assignment of these resonances.[3]

Later work in which separated isotope targets[4] were used and positron activity from both $Mg^{24}(p,\gamma)Al^{25}$ and $Mg^{25}(p,\gamma)Al^{26}$ were observed has been reported elsewhere. The γ-ray yield from a separated Mg^{25} target has also been observed and the resonances from magnesium have been identified as follows:

$Mg^{24}(p,\gamma)Al^{25}$: 225.5±0.2 and 418.4±0.5 kev.

$Mg^{25}(p,\gamma)Al^{26}$: 316.7±0.7, 391.5±0.5, 436.5±0.4, 495.6 ±0.6, 513.4±0.7, 530.4±0.7 kev.

$Mg^{26}(p,\gamma)Al^{27}$: 338.5±0.5 and 454.2±0.3 kev.

The positron yield from a natural magnesium target and the γ-ray yield from a separated Mg^{25} target as a function of proton energy are shown in Fig. 1.[5] The higher statistical accuracy obtained from the positron

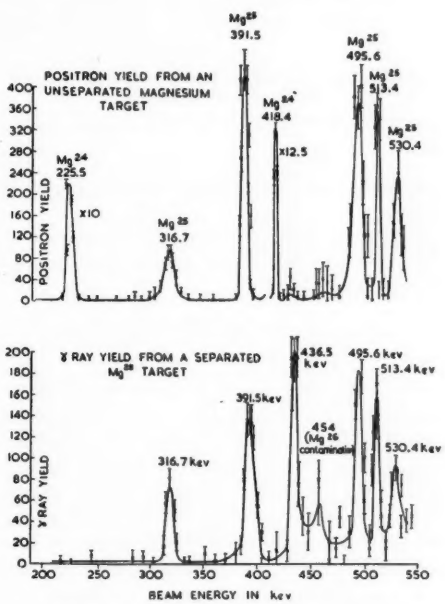


FIG. 1. Positron and γ-ray yield curves for the $Mg^{25}(p,\gamma)Al^{26}$ reaction.

yield curves has made possible more accurate estimations of these resonances, except in the case of the 436.5-kev resonance for which the positron yield was not high enough to be detected, though the γ-ray yield from this resonance was greater than that from the other $Mg^{25}(p,\gamma)Al^{26}$ resonances.

The resonance to which Kluyver refers as a doublet at 508 kev is seen from both the positron and γ-ray yield curves to consist of three distinct peaks in agreement with the results obtained by Tangen.[3]

Taking Kluyver's value of 6.350 Mev for the Q value of the proton capture process, the present results indi-

cate excited levels in Al^{26} at 6.655, 6.720, 6.769, 6.820, 6.840; and 6.860 Mev above the ground state.

Stähelin[6] and Moszkowski and Peaslee[7] have suggested that the level which decays by positron emission of 6.7-second half-life to Mg^{26} is not the ground state but the $T_z=0$ component of the lowest triplet with isobaric spin $T=1$ and ordinary spin $J=0^+$. The decay of the 6.769-Mev level by γ emission to this state therefore appears to be forbidden because of the absence of positron yield from the 436.5-kev resonance. This is consistent with Kluyver's observation of a strong 6.77 ± 0.08 Mev γ ray from the 6.769 level. The γ-ray de-excitation is presumed to be preferentially to the ground state of Al^{26}, which Kluyver postulates to have a spin of $J=5^+$. This in turn decays by a second forbidden positron transition to Mg^{26}. His estimated half-life of 4×10^4 years would insure that the positron yield was much too small to be detected in our observations.

It is of interest to note that the ratio of positron yield to γ-ray yield of the other resonances does not vary by more than a factor of 2 (Table I). If the probability of

TABLE I. Relative positron and γ-ray yields from the $Mg^{25}(p,\gamma)Al^{26*}$ resonances.

Resonant proton energy	γ-ray yield at peak	β^+ yield at peak	Ratio β^+/γ-ray yield
316.7	70	100	1.4
391.5	140	400	2.8
436.5	200	<20	<0.1
495.6	180	350	1.9
513.4	155	320	2.0
530.4	92	220	2.4

de-excitation of one of these levels to the 6.7 second β^+ activity level is Γ_1, and the total transition probability from the level is Γ_T, the above observation would seem to indicate that the ratio Γ_1/Γ_T is approximately constant for all the excited Al^{26} states listed here, except the 6.769 level. This conclusion may be modified slightly because of the possibility of de-excitation of some of these levels by cascade γ-ray emission as is indicated by Kluyver's detection of low-energy γ radiation.

Positrons from Al^{25} have been found to have an end-point energy approximately equal to those from Al^{26},[8] and the half-lives of the two activities are approximately equal.[9] It is therefore of interest to note that while the γ-ray yield of the 391.5-kev $Mg^{25}(p,\gamma)Al^{26}$ and 418.4-kev $Mg^{24}(p,\gamma)Al^{25}$ resonances are approximately equal,[2,3] the positron yield from the Mg^{24} resonances is approximately ten times higher than that from the Mg^{25} resonance (unseparated magnesium targets of comparable thickness were used for both measurements).

This could indicate that Γ_1/Γ_T for the excited Al^{26} levels is small compared with Γ_1/Γ_T for the Al^{25} levels. That is if we assume that Γ_1/Γ_T is unity for the Al^{25} level produced by the 418.4-kev resonance, an upper limit of $\frac{1}{10}$ is placed on the Γ_1/Γ_T ratio for the Al^{26} levels listed here. Alternatively, it could be postulated that there is a competing transition from the Al^{26} short-lived positron state itself, possible γ-ray transition to the ground state, which would also produce the short half-life of this state, and reduce the positron yield.

The authors would like to thank Dr. J. C. Kluyver, who initially drew our attention to the difference between his γ-ray yield curve and our positron yield curve from the $Mg^{25}(p,\gamma)Al^{26}$ reaction. The only remaining discrepancy between the two results is the existence of the resonance of 530.4 kev, which we find in both the positron and γ-ray yield curves.

They would also like to acknowledge helpful discussions with Professor J. Rotblat, Mr. D. R. Chick; and Mr. W. M. Jones. Thanks are also due to Dr. T. E. Allibone, F.R.S., for permission to publish the Letter.

[1] Kluyver, Van der Leun, and Endt, Phys. Rev. **94**, 1795 (1954).
[2] S. E. Hunt and W. M. Jones, Phys. Rev. **89**, 1283 (1953).
[3] R. Tangen; Kgl. Norske. Videnskab. Selskabs, Skrifter No. 1 (1946).
[4] Supplied by Dr. R. H. V. M. Dawton and Dr. M. L. Smith, Atomic Energy Research Establishment, Harwell.
[5] The positron yield curve is reproduced by kind permission of the Physical Society.
[6] P. Stähelin, Helv. Phys. Acta **26**, 691 (1953); Phys. Rev. **92**, 1076 (1953).
[7] S. A. Moszkowski and D. C. Peaslee, Phys. Rev. **93**, 455 (1954).
[8] Hunt, Jones, and Churchill, Proc. Phys. Soc. (London) **A68**, 479 (1954).
[9] Hunt, Jones, Churchill, and Hancock, Proc. Phys. Soc. (London) **A68**, 443 (1954).

Radiative Corrections to Muon Decay

R. J. FINKELSTEIN AND R. E. BEHRENDS
University of California, Los Angeles, California
(Received November 26, 1954)

THE instabilities of the neutron and the muon may be attributed to Fermi interactions of approximately the same strength. To compare these couplings in more detail, one must take into account the fact that these two processes are influenced differently by other fields: in particular, mesonic[1,2] and electromagnetic[3] corrections are present in neutron decay while only electromagnetic corrections are present in muon decay. In this Letter, the electromagnetic corrections to the spectrum of the muon are given for three typical interactions ($S+aP$, $S+aT+bP$, $S+aA+bP$, $a=\pm1$ and $b=\pm1$ in charge retention order). One might expect these corrections to be rather large since the energy of the electron near the end point is of the order of 100 rest masses. Further, without explicit calculation it is not possible to tell how seriously such an energy dependent correction distorts the observed spectrum. Since the nature of the interaction is inferred from this shape (especially near the end point, which is just where the radiative corrections are largest), it is important to consider this question carefully.

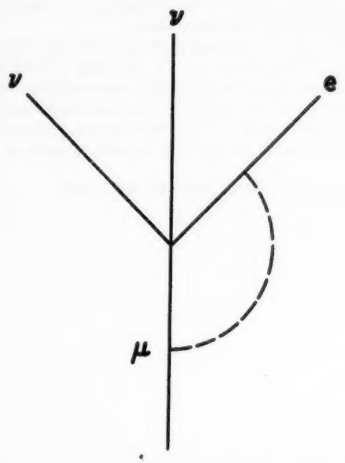


FIG. 1. A lowest order correction to μ decay.

The process is described by the familiar three diagrams, one of which is shown in Fig. 1. In addition one has to take into account inner bremsstrahlung,[4] which is of the same order. As usual the infinite infrared terms associated with the inner bremsstrahlung exactly cancel the corresponding infinite nonradiative terms. The really divergent terms associated with high photon energies cancel for the vector and axial vector interactions as in electrodynamics but in the other cases there are uncompensated logarithmically divergent terms.[5] The results, however, are rather insensitive to these terms: with a cutoff corresponding to the proton Compton wavelength, they amount to about one-fifth of the net correction.

These remarks are illustrated by the following expressions, valid for the scalar case, which express the change in the probability of decay associated with inner bremsstrahlung and virtual photons respectively.

$$\left(\frac{\Delta P}{P}\right)_{\text{I.B.}} = \frac{e^2}{2\pi}\left\{4\left(\ln\frac{2\epsilon}{m_2}-1\right)\ln\frac{m_1}{\lambda_{\min}}+4+\ln 2-\frac{2\epsilon_m}{\epsilon}\right.$$
$$+2\sum_{n=1}^{\infty}\frac{1}{n^2}\left(\frac{\epsilon}{\epsilon_m}\right)^n-2\left(\frac{\epsilon+\epsilon_m}{\epsilon}\right)\ln\frac{\epsilon_m'-\epsilon}{\epsilon_m}$$
$$\left.-\left[4\ln\frac{\epsilon_m'-\epsilon}{2\epsilon_m}+\frac{\epsilon_m^2-4\epsilon_m\epsilon-17\epsilon^2}{3\epsilon^2}\right]\ln\frac{m_2}{2\epsilon}\right\}, \quad (1)$$

$$\left(\frac{\Delta P}{P}\right)_{\text{V.}} = \frac{e^2}{2\pi}\left\{-4\left(\ln\frac{2\epsilon}{m_2}-1\right)\ln\frac{m_1}{\lambda_{\min}}\right.$$
$$+A+2\sum_{n=1}^{\infty}\frac{1}{n^2}\left(\frac{\epsilon}{\epsilon_m}\right)^n$$
$$\left.+\left[\frac{2\epsilon_m}{\epsilon_m'-\epsilon}+2\ln\frac{\epsilon_m'-\epsilon}{\epsilon}\right]\ln\frac{\epsilon}{\epsilon_m'}\right\}, \quad (2)$$

where

$$A=6\ln(\lambda/m_1)+\ln(m_1/m_2)-\tfrac{1}{2}$$
$$-\tfrac{1}{3}\pi^2+2[\ln(m_1/m_2)]^2=71.4,$$

$e^2=1/137$, $\epsilon_m'=\epsilon_m+m_2^2/2m_1$, ϵ = electron energy $=(p^2+m_2^2)^{\frac{1}{2}}$, ϵ_m = maximum electron energy $=m_1/2$, m_2 = mass of electron, m_1 = mass of muon, λ = maximum virtual photon energy, and $\lambda_{\min}$ = infrared cutoff.

The dominant term in (2) is $2[\ln(m_1/m_2)]^2$. The net correction is plotted in Fig. 2, which shows that $\Delta P/P$

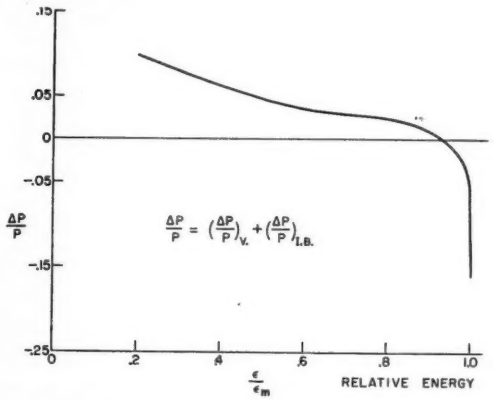


FIG. 2. Relative change in probability of decay for scalar case.

decreases from +7 percent at low energies to −17 percent at the endpoint. The downward trend of $\Delta P/P$ may be attributed to the fact that inner bremsstrahlung is relatively more important at low energies while virtual photons dominate at higher energies.

In Fig. 3, the results are given for $S+aP$, $S+aT+bP$, and $S+aA+bP$. The experimental points are those of Sagane *et al.*[6] The changes in the lifetime in the three cases are about 5 percent, 2.5 percent, and 3 percent respectively, which are rather large for electromagnetic corrections and must be taken into account in comparing

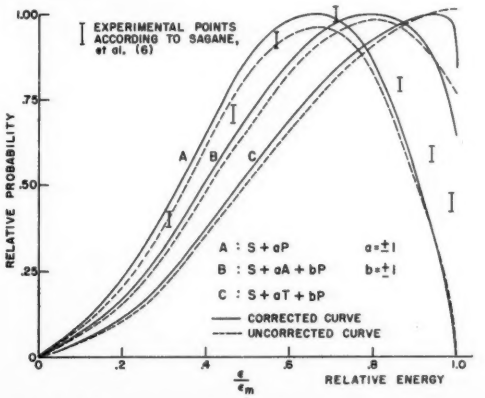


FIG. 3. Relative probabilities of decay with and without radiative corrections.

the neutron and muon decay. On the other hand, these corrections do not distort the shape of the spectrum in a significant way. Detailed calculations will be published.

[1] T. Kotani *et al.*, Progr. Theoret. Phys. (Japan) **6**, 1007 (1951).
[2] R. J. Finkelstein and S. A. Moszkowski, Phys. Rev. **95**, 1695 (1954).
[3] T. Nakano *et al.*, Progr. Theoret. Phys. (Japan) **5**, 1014 (1950).
[4] A. Lenard, Phys. Rev. **90**, 968 (1953). The differential cross sections given by Lenard were integrated to evaluate the total cross section of inner bremsstrahlung.
[5] S. Hanawa and T. Miyazima, Progr. Theoret. Phys. (Japan) **5**, 459 (1950); T. Nakano *et al.*, see reference 3. The first of these treats the correction to the decay of a meson with integral spin, the second to the decay of a neutron. The dominant term in the boson case is the same as in (2). The finite terms in the neutron case are not given. The same divergences appear in both of these calculations.
[6] R. Sagane *et al.*, Phys. Rev. **95**, 863 (1954).

CONTENTS—*Continued*

Letters to the Editor:

Quenching-in of Lattice Vacancies in Pure Gold . . . . . J. W. Kauffman and J. S. Koehler 555

Ferromagnetic Resonance in Nickel Ferrite Between One and Two Kilomegacycles . H. Suhl 555

Ultrasonic Attenuation Due to Lattice-Electron Interaction in Normal Conducting Metals . . . . . W. P. Mason 557

Can Helium-3 Be Expected to Exhibit Superfluidity at Sufficiently Low Temperatures? . . . . . O. K. Rice 558

Ionization and Solubility in Semiconductors . . . . . Howard Reiss and C. S. Fuller 559

K Band in Additively Colored KCl at 18°K . . . . . Felix E. Geiger, Jr. 560

Upper Limit for the Lifetimes of Excited States of Ni^{60} . . . . . Z. Bay, V. P. Henri, and F. McLernon 561

Inelastic Collision Cross Sections at 1.0-, 4.0-, and 4.5-Mev Neutron Energies . . . . . J. R. Beyster, R. L. Henkel, and R. A. Nobles 563

Multiplicity of Neutrons from the Spontaneous Fission of Californium-252 . . . . . Donald A. Hicks, John Ise, Jr., and Robert V. Pyle 564

Evidence for a Second Naturally Occurring Isotope in Tantalum . . . . . J. E. Evans, E. G. Joki, and R. R. Smith 565

New Naturally Occurring Isotope of Tantalum . . . . . F. A. White, T. L. Collins, Jr., and F. M. Rourke 566

Proton-Gamma Resonances in Magnesium . . . . . S. E. Hunt and D. A. Hancock 567

Radiative Corrections to Muon Decay . . . . . R. J. Finkelstein and R. E. Behrends 568

CONTENTS

THE PHYSICAL REVIEW SECOND SERIES, VOLUME 97, NO. 2 JANUARY 15, 1955

Departure from Paschen's Law of Breakdown in Gases . . . . . W. S. Boyle and P. Kisliuk 255
Nuclear Susceptibility of Liquid and Solid He^3 . . . . . P. J. Price 259
Interpretation of the Magnetic Behavior of Liquid Helium-3 . . . . . O. K. Rice 263
Breakdown Processes in Nitrogen, Oxygen, and Mixtures . . . . . Elsa L. Huber 267
Corona Threshold and the Range of Velocities of Pulse Spread in Geiger Counters . . . Leonard B. Loeb 275
F Centers in Pure and Hydride-Containing Alkali Halide Crystals . . . . R. S. Alger and R. D. Jordan 277
Structure in Optical Absorption of Barium Oxide Films . . . . . R. J. Zollweg 283
Dynamics of Simple Lattices . . . . . Herbert B. Rosenstock 289
Neutron Diffraction Studies of the Magnetic Structure of Alloys of Transition Elements . . . . . C. G. Shull and M. K. Wilkinson 304
Trapping of Minority Carriers in Silicon. I. *P*-Type Silicon . . . . . J. A. Hornbeck and J. R. Haynes 311
Recombination Processes in Insulators and Semiconductors . . . . . Albert Rose 322
Theory of the Faraday and Kerr Effects in Ferromagnetics . . . . . Petros N. Argyres 334
Neutron Damage to the Structure of Vitreous Silica . . . . . Joseph S. Lukesh 343
Photostimulated Emission of Some Activated Alkali Halide Phosphors . . . . . C. E. Mandeville and H. O. Albrecht 347
Ground State of Impurity Atoms in Semiconductors Having Anisotropic Energy Surfaces . . . . . Murray A. Lampert 352
Effect of Crystal Distortion upon Change of Resistivity of Silicon by Heat Treatment . . . . . W. C. Dash 354
Temperature Dependence of Electron Mobility in AgCl . . . . . Frederick C. Brown 355
Vibration Spectra and Specific Heats of Cubic Metals. I. Theory and Application to Sodium . A. B. Bhatia 363
Scattering of 0.6-, 1.0-, and 1.7-Mev Electrons from Aluminum and Gold . . . . . Robert T. Bayard and J. L. Yntema 372
Theory of Hyperfine Structure . . . . . Charles Schwartz 380
Shape of Collision-Broadened Spectral Lines . . . . . E. P. Gross 383
Charge Equilibrium Ratios for Hydrogen Ions from Solids . . . . . James A. Phillips 388
Magnetic Scattering of Slow Neutrons from O_2 Gas . . . . . W. H. Kleiner 411
Internal and External Bremsstrahlung Accompanying the Beta Rays of P^{32} . . . K. Lidén and N. Starfelt 419
Excitation Function for Na^{22} from Deuterons on Aluminum . . . . . S. O. Ring and L. M. Lidofsky 427
Energy Levels in Ca^{40} and Ca^{44} . . . . . John P. Schiffer 428
Beta Spectrum of C^{14} . . . . . A. V. Pohm, R. C. Waddell, J. P. Powers, and E. N. Jensen 432
Experiments on the Direct Photonuclear Effect . . . . . Sven A. E. Johansson 434
Effects of Isotopic Spin Selection Rules on Photonuclear Yields . . . . . H. Morinaga 446
Energy of the Ground State of Li^6 . . . . . J. Irving and D. S. Schonland 446
Radioactive Decay of Cs^{134} and Ce^{134m} . . . . . G. L. Keister, E. B. Lee, and F. H. Schmidt 451
Deuterium He^3 Reaction . . . . . W. E. Kunz 456
β Spectrum of Bi^{210} (RaE) and the Coupling Constants of Scalar and Tensor Interactions in β Decay . . . . . G. E. Lee-Whiting 463
Charged Photomesons from Carbon . . . . . David Luckey 469
Synchrotron Oscillations Induced by Radiation Fluctuations . . . . . Matthew Sands 470
Characteristics of Bubble Chambers . . . . . Donald A. Glaser and David C. Rahm 474
Scattering of 119-Mev Pions by Deuterium . . . . . D. E. Nagle 480
Landau Distribution and Density Effect at High Gas Pressures . . . . . E. D. Palmatier, J. T. Mears, and C. M. Askey 486
Theory of Multiple Coulomb Scattering from Extended Nuclei . . Leon N. Cooper and James Rainwater 492
Gauge Invariance and Renormalization in Electrodynamics . . . . . Dominique Rivier 505
Quantum-Mechanical Operators . . . . . V. Rojansky 507
Use of Causality Conditions in Quantum Theory . . . . . Marvin L. Goldberger 508
Geons . . . . . John Archibald Wheeler 511
Angular Momentum of a Real Field . . . . . R. H. Dicke 524
Goldstein's Eigenvalue Problem . . . . . H. S. Green 529
Coulomb Functions for Large Charges and Small Velocities . . . . . L. C. Biedenharn, R. L. Gluckstern, M. H. Hull, Jr., and G. Breit 542

(*Continued on cover three*)

LANCASTER PRESS, INC., LANCASTER, PA.